Amazon EMR 릴리즈 가이드

Amazon EMR

Amazon EMR: Amazon EMR 릴리즈 가이드

Table of Contents

Amazon EMR 릴리스에 대한 정보

Amazon EMR 릴리스는 빅 데이터 에코시스템의 오픈 소스 애플리케이션입니다. 각 릴리스는 클러스터를 생성할 때 Amazon EMR을 설치하고 구성하도록 선택한 여러 빅 데이터 애플리케이션, 구성 요소 및 기능으로 구성됩니다. 애플리케이션은 하둡 에코시스템과 관련된 오픈 소스 프로젝트 Apache BigTop 기반의 시스템을 사용하여 패키징됩니다. 이 안내서에서는 Amazon EMR 릴리스에 포함된 애플리케이션에 대한 정보를 제공합니다.

Amazon EMR을 시작하고 사용하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서를 참조하세요.

클러스터를 시작할 때 Amazon EMR의 여러 릴리스 중에서 선택할 수 있습니다. 이를 통해 호환성 요구 사항에 맞는 애플리케이션 버전을 테스트하고 사용할 수 있습니다. 릴리스 레이블과 함께 릴리스 번호를 지정합니다. 릴리스 레이블은 `emr-x.x.x` 형식입니다. 예를 들어 `emr-7.9.0`입니다.

Amazon EMR 5.18.0부터 Amazon EMR 아티팩트 리포지토리를 사용하여 특정 Amazon EMR 릴리스 버전에서 사용할 수 있는 정확한 버전의 라이브러리와 종속 항목에 대해 작업 코드를 빌드할 수 있습니다. 자세한 내용은 Amazon EMR 아티팩트 리포지토리를 사용하여 종속성 확인 단원을 참조하십시오.

새로운 Amazon EMR 릴리스가 출시될 때 업데이트를 받으려면 Amazon EMR 릴리스 정보에 대한 RSS 피드를 구독하세요.

Amazon EMR 7.x, 6.x 및 5.x 시리즈의 애플리케이션 버전, 릴리스 정보, 구성 요소 및 구성 분류를 포함한 최신 릴리스 세부 정보:

- Amazon EMR 릴리스 7.9.0
- Amazon EMR 릴리스 6.15.0
- Amazon EMR 릴리스 5.36.2

> **Note**
> 새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

최신 Amazon EMR 릴리스에 대한 릴리스 정보 및 모든 릴리스 기록:

- 새로운 기능
- Amazon EMR 릴리스 정보 아카이브

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록:

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

각 Amazon EMR 릴리스의 세부 정보 및 릴리스 시리즈별 차이(해당하는 경우):

- Amazon EMR 7.x 릴리스 버전
- Amazon EMR 6.x 릴리스 버전
- Amazon EMR 5.x 릴리스 버전
- Amazon EMR 4.x 릴리스 버전
- Amazon EMR 2.x 및 3.x AMI 버전

Amazon EMR 표준 지원

Amazon EMR 릴리스 이해

Amazon EMR 릴리스는 EC2, EKS 또는 Serverless 플랫폼에서 오픈 소스 애플리케이션을 실행하는 데 필요한 소프트웨어를 제공하는 도구입니다. Amazon EMR 릴리스는 런타임 환경, 코어 엔진 및 추가 항목과 같은 세 부분으로 구성됩니다.

- 런타임 환경 - 런타임 환경에는 클러스터 또는 컨테이너를 시작할 때 Amazon Machine Image(AMI) 또는 기타 컨테이너 이미지의 일부로 포함된 운영 체제가 포함됩니다(예: Amazon Linux 2023). 또한 Amazon Corretto JDK와 같은 언어 런타임과 이미지가 설치하는 기타 관련 도구도 포함됩니다.
- 코어 엔진 - 여기에는 Apache Hive와 같은 코어 오픈 소스 소프트웨어와 Apache Iceberg와 같은 오픈 테이블 형식이 포함됩니다.
- 추가 항목 - 여기에는 Apache Pig와 같은 오픈 소스 소프트웨어, `mariadb-connector-java`와 같은 Python 패키지, 편의 라이브러리가 포함됩니다.

선택한 Amazon EMR 릴리스 버전은 이러한 모든 구성 요소를 새 버전의 Amazon Linux(AL) 기반 Amazon Machine Images(AMI) 또는 컨테이너 이미지로 번들링합니다. Amazon EMR은 클러스터, EKS 컨테이너 또는 서버리스 애플리케이션의 모든 노드에 대해 동일한 AL 버전을 유지 관리합니다. Amazon EMR은 Amazon Linux에서 릴리스한 후 90일 이내에 최신 런타임 환경을 가져오고 업스트림 릴리스 후 90일 이내에 새로운 오픈 소스 버전의 코어 엔진을 확보하는 것을 목표로 합니다. 필요에 따라 추가 항목이 릴리스됩니다. 각 Amazon EMR 릴리스에 대한 릴리스 정보에서 소프트웨어 목록을 찾을 수 있습니다.

릴리스 버전 관리

각 릴리스의 변경 범위를 이해하는 데 도움이 되도록 Amazon EMR은 시멘틱 버전 관리를 사용합니다. 시맨틱 버전 관리는 `<major>.<minor>.<patch>` 형식을 따릅니다. 예를 들어 7.0.0 릴리스를 고려합니다.

첫 번째 자리는 메이저 버전 릴리스 또는 메이저 릴리스를 나타냅니다. 메이저 릴리스는 일반적으로 역호환되지 않는 런타임 환경 또는 코어 엔진에 대한 상당한 변경 사항, 개선 사항 및 새로운 기능을 제공합니다. 예를 들어 7.x 메이저 릴리스는 기본값으로 Amazon Corretto 17 JDK와 함께 Amazon Linux 2023을 사용합니다. 따라서 릴리스에는 기본값으로 Amazon Corretto 8 JDK를 사용하고 Amazon Linux 2에서 제공하는 6.x 메이저 릴리스와 비교했을 때 몇 가지 혁신적인 변경 사항이 포함되어 있습니다.

두 번째 숫자는 마이너 버전 릴리스 또는 마이너 릴리스를 나타냅니다. 마이너 릴리스는 코어 엔진 및 추가 항목, 새 기능에 대한 증분 변경 사항, 개선 사항 및 기능을 포함하는 역호환되지 않는 릴리스입니다. 예를 들어 6.15는 6.11.1과 비교했을 때 Apache Spark 3.3.2와 함께 제공되는 Apache Spark 3.4.1을 사용합니다.

세 번째 숫자는 패치 버전 릴리스 또는 패치 릴리스를 나타냅니다. 패치 릴리스는 코어 엔진 및 추가 항목에 대한 수정 사항 및 패치 업데이트를 포함하는 역호환되는 릴리스이지만, 새로운 기능이나 OSS 버전 업그레이드를 포함하지 않습니다. 예를 들어 6.11.1 및 6.11.0 모두 동일한 Apache Spark 3.3.2 OSS 버전을 포함합니다. 6.6.x부터 최신 패치 릴리스는 사용 가능한 최신 런타임 환경 패치를 제공합니다. 런타임 환경은 Amazon EMR의 패치 릴리스에 영향을 주지 않습니다. 예를 들어 6.11.1은 2023년 12월 19일부터 2024년 1월 10일까지 AL 버전 2.0.20240109.0으로 시작되었지만, 2024년 1월 11일부터 AL 버전 2.0.20240124.0으로 시작됩니다. 일부 Amazon EMR on EKS 릴리스를 사용하면 시맨틱 버전에 태그 또는 접미사를 추가하여 대체 기능을 제공할 수 있습니다. 예를 들어, Amazon EMR on EKS의 `emr-6.15.0-latest` 릴리스는 Amazon Corretto 8 JDK 지원 릴리스 버전을 시작하는 반면, Amazon EMR on EKS의 `emr-6.15.0-java17-latest` 릴리스는 Amazon Corretto 17 JDK 지원 릴리스 버전을 시작할 수 있습니다.

지원 정책

소개

Amazon EMR은 최소 90일마다 1회 새 마이너 릴리스를 발행하고 릴리스 날짜로부터 24개월 동안 마이너 릴리스에 대한 지원을 제공하는 것을 목표로 합니다. 이 지원은 런타임 환경 및 코어 엔진과 관련 종속 항목을 포함합니다. 단, 추가 항목 또는 관련 종속 항목은 포함하지 않습니다. 런타임 환경에는 클러스터 및 애플리케이션을 시작, 중지 및 운영하는 데 필요한 독점 플랫폼 구성 요소 및 API도 포함되어 있습니다. 이 일관된 릴리스 일정은 예측 가능한 주기를 보장하므로 지원 및 보안을 제공하는 버전을 보다 편리하게 계획 및 테스트하고 해당 버전으로 전환할 수 있습니다.

표준 지원에서 기대할 수 있는 사항

표준 지원은 권장 구성에 따라 런타임 환경 및 코어 엔진 구성 요소에서 발생하는 문제에 대한 기술 지원 티켓에서 수정 사항을 제공합니다. 모든 수정 사항은 가용성에 따라 달라질 수 있습니다. 권장 구성은 Amazon EMR 설명서에서 제공하는 경우를 제외하고 Amazon EMR 릴리스에 있는 바이너리 및 구성에 대한 수정, 추가 또는 변경 없이 Amazon EMR을 사용하는 것입니다. Amazon EMR은 수정 사항을 확인한 후 90일 이내에 Amazon EMR 릴리스의 최신 패치, 마이너 또는 메이저 버전에 수정 사항을 배포합니다. Amazon EMR은 EC2 클러스터에서 새 EMR을 시작하거나, 새 Amazon EMR on EKS 컨테이너를 시작하거나, 새 EMR Serverless 작업을 트리거할 때 자동으로 수정 사항을 적용합니다. 추가 항목 구성 요소는 편의를 위해 제공되며 Amazon EMR은 추가 항목과 관련된 수정 사항을 제공하지 않습니다.

Components

표준 지원에는 운영 체제, 언어 런타임, Apache Hive 및 Apache Iceberg와 같은 코어 오픈 소스 소프트웨어 등의 런타임 환경 및 코어 엔진 구성 요소가 포함됩니다. Amazon EMR 릴리스에 대한 정보 및 Amazon Linux FAQ에서 각 릴리스에 대해 지원되는 구성 요소의 전체 목록을 확인할 수 있습니다.

다음 목록은 표준 지원에서 다양한 구성 요소 유형에 대해 제공하는 지원을 설명합니다.

- 런타임 환경 구성 요소: 런타임 환경 구성 요소는 기술 지원 티켓에 대한 수정 사항을 수신합니다. 수정 사항은 (a) 중요한 버그, (b) 중요한 데이터 손상 문제, (c) 중요한 보안 문제로 분류됩니다. 적합한 경우 Amazon EMR은 런타임 환경에 대한 수정 사항을 이전 버전으로 백포트합니다. 오픈 소스 코어 엔진 구성 요소와의 운영 호환성을 위해 특정 런타임 환경 구성 요소는 고객 애플리케이션 중단을 방지하고자 특정 버전에 남아 있어야 합니다. 이러한 구성 요소의 경우 Amazon EMR은 수정 사항 가용성을 위해 업스트림 오픈 소스에 의존합니다. 오픈 소스에서 수정 사항을 사용할 수 있는 경우 Amazon EMR에서 수정 사항을 확인한 후 90일 이내에 안정적인 최신 버전을 제공합니다.
- 코어 엔진 구성 요소: 코어 엔진은 각각 수백 개의 전이적 종속 라이브러리를 보유한 많은 오픈 소스 프로젝트에 대한 최신 버전을 제공합니다. 이러한 프로젝트를 관리하는 오픈 소스 커뮤니티는 문제

와 알려진 일반적인 취약성 및 노출(CVE)을 해결하려고 자주 시도하지만 최신 버전에는 여전히 알려진 버그 및 CVE가 포함될 수 있습니다. Amazon EMR은 업스트림 오픈 소스의 수정 사항 가용성에 종속되어 있으며, Amazon EMR에서 확인한 후 90일 이내에 코어 엔진 구성 요소의 일부로 안정적인 최신 버전을 제공합니다. 경우에 따라 Amazon EMR은 코어 엔진 구성 요소 중 하나에 업스트림 오픈 소스보다 먼저 해결해야 하는 CVE에 대한 수정 사항을 제공할 수 있습니다. 또한 Amazon EMR은 권장 구성에 따라 오픈 소스에 추가된 기능에 대한 기술 지원 및 수정 사항도 제공합니다. 코어 엔진 구성 요소에 대한 수정 사항을 이전 패치 또는 마이너 버전으로 백포트하지 않습니다.

- 추가 구성 요소: Amazon EMR은 추가 항목 구성 요소를 지원하지 않습니다. 추가 항목 구성 요소는 편의상 제공되는 오픈 소스 프로젝트이며 Amazon EMR은 이러한 구성 요소에서 발생하는 문제에 대한 수정 사항을 제공하지 않습니다. 모든 지원 요청 또는 수정 사항은 이러한 구성 요소를 지원하는 오픈 소스 커뮤니티를 통해 해결할 수 있습니다.

표준 지원 수명 주기

다음에서는 표준 지원 수명 주기의 마일스톤을 설명합니다.

- 표준 지원: Amazon EMR 릴리스는 릴리스 날짜로부터 24개월 이내에 표준 지원을 받을 자격이 주어집니다. 기술 지원 티켓을 생성하고 이러한 릴리스에서 발생하는 문제에 대한 업데이트를 예상할 수 있습니다.
- 지원 종료: 표준 지원이 종료된 후 Amazon EMR 릴리스는 12개월 동안 지원 종료(EoS) 단계에 들어갑니다. EoS 릴리스는 기술 지원을 받을 수 없으며 이러한 릴리스에서 실행되는 클러스터, 컨테이너 또는 작업에 대한 티켓을 생성할 수 없습니다. EoS 릴리스는 수정 사항, 패치 또는 업데이트를 수신하지 않습니다. EoS 릴리스는 콘솔에서 제거되지만, API 및 AWS CLI를 통해 계속 사용할 수 있습니다. 언제든지 EoS 릴리스에서 워크로드를 계속 실행할 수 있습니다. 최신 Amazon EMR 릴리스로 마이그레이션하여 보안 패치를 계속 받고, 기술 지원 자격을 유지하며, 필요한 경우 지원 티켓을 생성할 수 있도록 하는 것이 좋습니다.
- 수명 종료: 지원 종료(EoS) 기간 이후 릴리스는 수명 종료(EoL)로 간주됩니다. EoL 클러스터를 계속 실행할 수 있지만 Amazon EMR은 보안 및 운영 문제로 인해 사례별로 API 및 SDK에서 EoL 릴리스를 제거할 수 있는 권한을 보유합니다. 예외적인 경우 API 및 SDK에서 EoL 버전을 제거할 수 있으므로 최신 Amazon EMR 릴리스로 마이그레이션하는 것이 좋습니다.

브리지 지원

Amazon EMR은 2024년 7월 25일에 이 새로운 지원 정책을 발표했습니다. 이 정책에 따라 2022년 7월 24일 또는 그 이전에 릴리스된 Amazon EMR 버전이 이제 지원 종료 대상으로 지정됩니다. 그러나 최

신 버전을 계획하고 마이그레이션할 수 있는 추가 시간을 제공하기 위해 Amazon EMR은 이 발표 전 2년 이내에 릴리스된 해당 이전 버전에 대해 표준 지원과 동등한 브리지 지원을 제공합니다.

2024년 7월 25일 이후에는 릴리스 정보에서 현재 상태 및 지원 일정을 참조할 수 있습니다.

다음 표는 2024년 7월 25일 정책 발표 당시, 기존 Amazon EMR 릴리스에 대한 지원 상태를 보여줍니다.

릴리스 및 지원 기간

Amazon EMR 릴리스 버전	초기 릴리스 날짜	표준 지원 종료 날짜	표준 종료 시작 날짜	수명 종료 시작 날짜
7.2.0	2024년 7월 25일	2026년 7월 24일	2026년 7월 25일	2027년 7월 25일
7.1.0	2024년 4월 23일	2026년 4월 22일	2026년 4월 23일	2027년 4월 23일
7.0.0	2023년 12월 19일	2025년 12월 18일	2025년 12월 19일	2026년 12월 19일
5.36.x 및 6.6.x~6.15.x	2022년 5월 9일 ~2023년 11월 13일	2026년 1월 24일까지 브리지 지원	2026년 1월 25일	2026년 1월 25일
• 6.x 시리즈: 6.5.0 이하 • 5.x 시리즈: 5.35.0 이하 • 4.x, 3.x 및 2.x 시리즈	2013년 1월 31일 ~2022년 3월 30일	2025년 7월 24일까지 브리지 지원	2025년 7월 25일	2025년 7월 25일

고려 사항

표준 지원은 Amazon EMR을 사용할 수 있는 모든 리전의 모든 Amazon EMR 배포 모델(EMR on EC2, Amazon EMR on EKS, EMR Serverless)에 대해 추가 비용 없이 제공됩니다. 권장 구성으로 실행되는 클러스터는 정책에 설명된 대로 자동으로 지원을 받을 수 있으므로, 지원을 활성화하기 위해 추가 작업을 수행하지 않아도 됩니다.

- 표준 지원에서는 Amazon EMR 클러스터에 필요한 구성 요소만 지원합니다. Amazon EMR은 코어 엔진 오픈 소스 구성 요소가 EoL 업스트림에 도달하거나 종속 항목에 대한 보안 업데이트를 더 이상 사용할 수 없는 경우 보안 패치 및 수정 사항 가용성을 보장할 수 없습니다. 추가 항목 설치를 옵트인할 수 있지만 Amazon EMR은 추가 항목 또는 해당 종속 항목을 지원하지 않습니다. 예를 들어 사용자 지정 AMI에서 서드파티 애플리케이션을 설치하여 클러스터의 보안을 강화하거나, 부트스트랩 작업 스크립트를 사용하여 추가 구성 요소를 설치하거나 객체를 복사하거나, 클러스터에 SSH로 연결하여 기본 패키지 버전을 업그레이드할 수 있습니다. Amazon EMR은 이러한 구성 요소를 지원하지 않습니다. 표준 지원은 고객이 제공한 부트스트랩 작업, 패키지, 라이브러리, 사용자 지정 코드 및 사용자의 편의를 위해 설치하도록 Amazon EMR을 구성할 수 있는 기존 보유 사용자 지정 애플리케이션에 적용되지 않습니다.
- 실행 중인 Amazon EMR 릴리스에 관계없이 기존 클러스터는 영향을 받지 않습니다. 중단 없이 기존 클러스터를 계속 실행할 수 있습니다. 또한 새 클러스터를 계속 시작하고 기존 릴리스 및 새 릴리스에서 작업을 실행할 수 있습니다. 정책이 발효되는 시점의 모든 기존 릴리스 및 새 릴리스는 Amazon EMR 릴리스 버전의 최초 릴리스 날짜로부터 24개월 동안 표준 지원의 적용을 받습니다. Amazon EMR은 정책의 초기 발표 중에 브리지 지원을 제공합니다. 중단 없는 지원을 받으려면 애플리케이션을 테스트하고 최신 Amazon EMR 릴리스로 즉시 업그레이드하는 것이 좋습니다.
- Amazon EMR은 기존 릴리스 또는 클러스터의 표준 지원 구성 요소를 변경하지 않습니다. 그러나 Amazon EMR은 사례별로 업스트림 수명 종료를 고려하고, 새 릴리스에서 이러한 구성 요소를 제거하거나, 예외적인 경우 기존 릴리스를 지원 종료(EoS) 또는 수명 종료(EoL) 상태로 전환할 수 있는 권한을 보유합니다. 사용 가능한 채널을 통해 제거 사항이 있으면 알려 드립니다.
- 새 리전이 시작되면 Amazon EMR은 표준 지원에 따라 Amazon EMR 릴리스 버전만 지원하며, 새 리전이 정식 출시되는 날짜로부터 6개월 이내에 릴리스됩니다.
- Amazon EMR은 기존 클러스터를 최신 버전으로 자동 업데이트하지 않습니다. 그러나 원하는 경우 새 클러스터를 최신 패치 버전으로 업데이트하도록 선택할 수 있습니다.
- 지원되지 않는 Amazon EMR 사용 비용은 AWS 청구서에 포함됩니다. 지원되지 않는 방식으로 Amazon EMR을 사용하더라도 해당 사용과 관련된 비용은 여전히 전체 AWS 소비의 일부이며 지원 요금 계산에 포함됩니다.

자세한 내용은 [AWS 개발자 지원]에 문의하세요.

Amazon EMR 7.x 릴리스 버전

이 섹션에는 각 Amazon EMR 7.x 릴리스 버전에 제공되는 애플리케이션 버전, 릴리스 정보, 구성 요소 버전 및 구성 분류가 포함되어 있습니다.

클러스터를 시작할 때 Amazon EMR의 여러 릴리스 중에서 선택할 수 있습니다. 이를 통해 호환성 요구 사항에 맞는 애플리케이션 버전을 테스트하고 사용할 수 있습니다. 릴리스 레이블과 함께 릴리스 번호를 지정합니다. 릴리스 레이블은 `emr-x.x.x` 형식입니다. 예를 들어 `emr-7.9.0`입니다.

새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

모든 Amazon EMR 7.x 릴리스에서 애플리케이션 버전의 전체 테이블은 Amazon EMR 7.x 릴리스의 애플리케이션 버전 섹션을 참조하세요.

주제

Amazon EMR 7.x 릴리스의 애플리케이션 버전

각 Amazon EMR 7.x 릴리스에서 사용할 수 있는 애플리케이션 버전을 표시한 전체 테이블을 보려면 브라우저에서 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 엽니다.

Amazon EMR 릴리스 7.9.0

7.9.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.9.0에 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜
초기 릴리스 날짜	2025년 5월 19일
표준 지원 종료 시점	2027년 5월 19일
지원 종료	2027년 5월 20일
수명 종료	2028년 5월 19일

7.9.0 릴리스 구성 요소

Amazon EMR 7.9.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta, Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, Oozie, Pig, Presto, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.9.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hive, Hadoop, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Presto, Pig, Spark, TensorFlow, Tez, Trino및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-7.9.0	emr-7.8.0	emr-7.7.0	emr-7.6.0
AWS Java용 SDK	2.31.16, 1.12.782	2.29.52, 1.12.780	2.29.44, 1.12.780	2.29.25, 1.12.779
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.3.0-amzn-1	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1
Flink	1.20.0	1.20.0	1.20.0	1.20.0
HBase	2.6.2	2.6.1	2.6.1	2.6.1
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.1	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-6	0.15.0-amzn-5	0.15.0-amzn-4	0.15.0-amzn-3
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-2	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0

	emr-7.9.0	emr-7.8.0	emr-7.7.0	emr-7.6.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.1	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.287	0.287	0.287	0.287
Spark	3.5.5	3.5.4	3.5.3	3.5.3
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	467	467	457
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.3	3.9.2	3.9.2

7.9.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.9.0에 대한 정보가 포함되어 있습니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.
- 영구 앱 UI에 대한 CM-CMK 지원 -이 기능은 EMR 서비스 버킷에서 관리되는 로그에 적용되는 SSE 암호화를 확장하여 고객 관리형 사용자 지정 마스터 키(CM-CMK)를 준수합니다. 이 기능을 사용하면 클러스터가 종료된 후 고객이 로그의 접근성을 관리할 수 있습니다.

알려진 문제

- EC2의 EMR에 대한 로깅 옵션이 비활성화된 경우 클러스터가 종료되고 영구 애플리케이션 UIs가 포함되지 않은 후에는 로그를 사용할 수 없습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태ExecutorState라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 TRUE (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Oozie 사용 중단 - OSS Oozie가 Attic 모드로 전환되었으므로 Apache Oozie를 포함하는 마지막 릴리스는 Amazon EMR 릴리스 7.9입니다. Apache Oozie는 Amazon EMR 릴리스 7.10 이상 릴리스에서 제거됩니다. Oozie 대신 MWAA를 사용하는 것이 좋습니다. 고객은 GitHub의 유틸리티 oozie-to-airflow-emr를 참조하여 기존 Oozie 워크플로를 MWAA로 마이그레이션할 수 있습니다.
- YARN 컨테이너 빈 패킹 - Amazon EMR 버전 7.9.0부터 YARN의 다중 노드 배치 정책을 기반으로 구축된 YARN 용량 스케줄러에 컨테이너 빈 패킹 정책을 사용할 수 있습니다.
- hudi-cli.sh - hudi-cli.sh 스크립트는 EMR 릴리스 7.9.0 이상에서 더 이상 사용되지 않습니다. Amazon EMR 릴리스 7.9.0 이상에서는 hudi-cli-bundle.jar을 사용합니다.
- FIPS 지원 AL2023 AMI의 EMR 클러스터 - EMR 7.9부터 고객은 FIPS 지원 AL2023 AMI에서 EMR 클러스터를 시작할 수 있습니다. 이 변경 사항은 모든 EMR 패키지가 FIPS 140-3에 대한 권장 암호화 정책을 준수하도록 EMR의 디지털 서명 및 해시 알고리즘을 업데이트합니다. 참고: EMR 또는 AL2023의 FIPS 인증 상태는 여기에 포함되지 않습니다. FIPS 140-3 검증 상태에 대한 최신 정보는 AL2023 FIPS FAQ를 참조하세요.
- 단계 지연 시간 -이 릴리스에서는 설치할 하둡 애플리케이션을 명시적으로 선택하지 않고 전송 중 데이터 암호화를 사용할 때 단계 지연 시간이 증가하는 문제를 해결합니다.
- EBS 암호화 볼륨을 사용한 클러스터 시작 -이 릴리스에서는 EBS 암호화 볼륨에 대한 KMS 호출로 인해 EMR 클러스터 시작이 실패할 수 있는 문제를 해결합니다.

•

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2일	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

7.9.0 기본 Java 버전

Amazon EMR 릴리스 7.9 이상은 Apache Livy를 제외하고 Corretto 17(OpenJDK)을 지원하는 애플리케이션에 기본적으로 Amazon Corretto 17(OpenJDK 기반)과 함께 제공됩니다.

다음 표에는 Amazon EMR 7.9.0의 애플리케이션에 대한 기본 Java 버전이 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

7.9.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.3.0-amzn-1	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.18.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.22.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.18.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.39.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.18.0	EMR S3Select 커넥터
emr-wal-cli	1.3.4	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.70.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.20.0-amzn-3	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.20.0-amzn-3	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.4.1-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.4.1-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.4.1-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.4.1-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.4.1-amzn-1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.4.1-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.4.1-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.4.1-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.4.1-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.4.1-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.4.1-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.6.2-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.6.2-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.6.2-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.6.2-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.6.2-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.6.2-amzn-0	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-18	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-18	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-18	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-18	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-18	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-18	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-18	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.15.0-amzn-6	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.15.0-amzn-6	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.15.0-amzn-6	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.15.0-amzn-6	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.7.1-amzn-2	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	12.3.0	Nvidia 드라이버와 Cuda 도구 키트

구성 요소	버전	설명
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.2.1	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.287-amzn-3	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.287-amzn-3	여러 쿼리를 실행하는 서비스.
presto-client	0.287-amzn-3	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	467-amzn-2	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	467-amzn-2	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
trino-client	467-amzn-2	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.5-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.5.5-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.5-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.5-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	25.02.1-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
tensorflow	2.16.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-16	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-16	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
zeppelin-server	0.11.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.3-amzn-1	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.3-amzn-1	ZooKeeper 명령줄 클라이언트.

7.9.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-7.9.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j2	HBase의 hbase-log4j2.properties 파일에서 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j2	Phoenix의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.

분류	설명	재구성 작업
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.

분류	설명	재구성 작업
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatc hAgent service.

7.9.0 변경 로그

7.9.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2025-05-19	문서 게시	Amazon EMR 7.9.0 릴리스 정보 최초 게시
2025-05-19	배포 완료	지원되는 모든 리전에 Amazon EMR 7.9.0 완전 배포
2025-05-19	초기 제한 릴리스	Amazon EMR 7.9.0이 초기 상용 리전에 처음 배포됨

Amazon EMR 릴리스 7.8.0

7.8.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.8.0에 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜	
초기 릴리스 날짜	2025년 3월 7일	

지원 단계	날짜
표준 지원 종료 시점	2027년 3월 6일
지원 종료	2027년 3월 7일
수명 종료	2028년 3월 7일

7.8.0 릴리스 구성 요소

Amazon EMR 7.8.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta, Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, Oozie, Pig, Presto, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.8.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, , Phoenix, Pig, Spark, PrestoTensorFlow, Tez, Trino및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.8.0	emr-7.7.0	emr-7.6.0	emr-7.5.0
AWS Java용 SDK	2.29.52, 1.12.780	2.29.44, 1.12.780	2.29.25, 1.12.779	2.28.8, 1.12.772
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.3.0-amzn-0	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0
Flink	1.20.0	1.20.0	1.20.0	1.19.1
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.6.1	2.5.10
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.1	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-5	0.15.0-amzn-4	0.15.0-amzn-3	0.15.0-amzn-2
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-1	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-

	emr-7.8.0	emr-7.7.0	emr-7.6.0	emr-7.5.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.1	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.287	0.287	0.287	0.287
Spark	3.5.4	3.5.3	3.5.3	3.5.2
Sqoop	-	-	-	-
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	467	457	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.3	3.9.2	3.9.2	3.9.2

7.8.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.8.0에 대한 정보가 포함되어 있습니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.8.0 application upgrades include AWS SDK for Java 2.29.52/1.12.780, Delta 3.3.0-amzn-0, Hadoop 3.4.1, Hudi 0.15.0-amzn-5, Iceberg 1.7.1-amzn-1, Spark 3.5.4, and Zookeeper 3.9.3.

알려진 문제

- 이 릴리스에는 적용되지 않습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.

•

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1일	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

7.8.0 기본 Java 버전

Amazon EMR 릴리스 7.8 이상은 Apache Livy를 제외하고 Corretto 17(OpenJDK)을 지원하는 애플리케이션에 기본적으로 Amazon Corretto 17(OpenJDK 기반)과 함께 제공됩니다.

다음 표에는 Amazon EMR 7.8.0의 애플리케이션에 대한 기본 Java 버전이 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

7.8.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.3.0-amzn-0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.17.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.21.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.18.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.38.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.17.0	EMR S3Select 커넥터
emr-wal-cli	1.3.3	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.69.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.20.0-amzn-2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.20.0-amzn-2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.4.1-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.4.1-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.4.1-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.4.1-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.4.1-amzn-0	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.4.1-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.4.1-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.4.1-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.4.1-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.4.1-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.4.1-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.6.1-amzn-2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.6.1-amzn-2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.6.1-amzn-2	HBase 명령줄 클라이언트
hbase-rest-server	2.6.1-amzn-2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.6.1-amzn-2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.6.1-amzn-2	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-17	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-17	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-17	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-17	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-17	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-17	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-17	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.15.0-amzn-5	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.15.0-amzn-5	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.15.0-amzn-5	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.15.0-amzn-5	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.7.1-amzn-1	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	12.3.0	Nvidia 드라이버와 Cuda 도구 키트

구성 요소	버전	설명
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.2.1	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.287-amzn-2	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.287-amzn-2	여러 쿼리를 실행하는 서비스.
presto-client	0.287-amzn-2	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	467-amzn-1	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	467-amzn-1	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
trino-client	467-amzn-1	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.4-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.5.4-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.4-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.4-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	24.12.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
tensorflow	2.16.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-15	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-15	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
zeppelin-server	0.11.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.3-amzn-0	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.3-amzn-0	ZooKeeper 명령줄 클라이언트.

7.8.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-7.8.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hb ase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j2	Phoenix의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.

분류	설명	재구성 작업
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.

분류	설명	재구성 작업
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatchAgent service.

7.8.0 변경 로그

7.8.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2025-03-07	문서 게시	Amazon EMR 7.8.0 릴리스 정보가 처음 게시됨
2025-03-06	배포 완료	지원되는 모든 리전에 Amazon EMR 7.8.0 완전 배포
2025-02-28	초기 제한 릴리스	Amazon EMR 7.8.0이 초기 상용 리전에 처음 배포됨

Amazon EMR 릴리스 7.7.0

7.7.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.7.0에 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜	
초기 릴리스 날짜	2025년 2월 6일	

지원 단계	날짜
표준 지원 종료 시점	2027년 2월 5일
지원 종료	2027년 2월 6일
수명 종료	2028년 2월 6일

7.7.0 릴리스 구성 요소

Amazon EMR 7.7.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta, Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, Oozie, Pig, Presto, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.7.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hive, Hadoop, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Presto, Pig, Spark, TensorFlow, Tez, Trino및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
AWS Java용 SDK	2.29.44, 1.12.780	2.29.25, 1.12.779	2.28.8, 1.12.772	2.28.8, 1.12.772
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.1-amzn-2	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1
Flink	1.20.0	1.20.0	1.19.1	1.19.1
Ganglia	-	-	-	-
HBase	2.6.1	2.6.1	2.5.10	2.5.5
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.4.0
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-4	0.15.0-amzn-3	0.15.0-amzn-2	0.15.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.7.1-amzn-0	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-

	emr-7.7.0	emr-7.6.0	emr-7.5.0	emr-7.4.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.1	5.2.0	5.2.0	5.2.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.287	0.287	0.287	0.287
Spark	3.5.3	3.5.3	3.5.2	3.5.2
Sqoop	-	-	-	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	467	457	446	446
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.2

7.7.0 릴리스 구성 요소 세부 정보

Amazon EMR 7.7.0의 추가 패키지에 대한 자세한 내용은 다음 표를 참조하세요.

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	추가 항목	3.4.0-amzn-3-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	s3-dist-cp	추가 항목	2.37.0-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper	추가 항목	3.9.2.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	추가 항목	3.9.2.amzn.1-1.amzn2023
Flink	Core	hadoop-client	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-httpfs	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-kms	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.3-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hadoop-yarn	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Flink	Core	hudi	추가 항목	0.15.0.amzn.4-1.amzn2023
Flink	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Flink	Core	hadoop	추가 항목	3.4.0.amzn.3-1.amzn2023
HBase	Core	s3-dist-cp	추가 항목	2.37.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
HCatalog	Core	mariadb-connector-java	추가 항목	

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Hive	Core	hadoop-httpfs	추가 항목	3.4.0.amzn.3-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Hive	Core	hive-hbase	추가 항목	3.1.3.amzn.16-1.amzn2023
Hive	Core	mariadb-connector-java	추가 항목	2.7.2-1
Hive	Core	s3-dist-cp	추가 항목	2.37.0-1.amzn2023
피닉스	Core	s3-dist-cp	추가 항목	2.37.0-1.amzn2023
Spark	Core	spark-datanucleus	추가 항목	3.5.3.amzn.1-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.3-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	bigtop-utils	추가 항목	1.2.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-client	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-hdfs	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-kms	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.3-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-yarn-resourcemanager	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Trino	Core	hive	추가 항목	3.1.3.amzn.16-1.amzn2023
Trino	Core	hive-hcatalog-server	추가 항목	3.1.3.amzn.16-1.amzn2023
Trino	Core	mariadb-connector-java	추가 항목	2.7.2-1
Trino	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Trino	Core	zookeeper	추가 항목	3.9.2.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	추가 항목	3.9.2.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	추가 항목	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	추가 항목	1.2.4-1.amzn2023
Trino	Core	hadoop	추가 항목	3.4.0.amzn.3-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-lzo	추가 항목	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	추가 항목	3.1.3.amzn.16-1.amzn2023
Trino	Core	hive-jdbc	추가 항목	3.1.3.amzn.16-1.amzn2023
Zookeeper	Core	emrfs	추가 항목	3.9.2.amzn.1-1.amzn2023
Zookeeper	Core	hadoop-client	추가 항목	3.9.2.amzn.1-1.amzn2023
Zookeeper	Core	hadoop-hdfs	추가 항목	3.9.2.amzn.1-1.amzn2023
Zookeeper	Core	hadoop-hdfs-datanode	추가 항목	3.9.2.amzn.1-1.amzn2023
Zookeeper	Core	hadoop-hdfs-journalnode	추가 항목	3.9.2.amzn.1-1.amzn2023
Zookeeper	Core	hadoop-hdfs-namenode	추가 항목	3.9.2.amzn.1-1.amzn2023
Zookeeper	Core	hadoop-httpfs	추가 항목	3.9.2.amzn.1-1.amzn2023
Zookeeper	Core	hadoop-kms	추가 항목	3.9.2.amzn.1-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.3-1.amzn2023
Zookeeper	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Zookeeper	Core	hadoop-yarn	추가 항목	3.4.0.amzn.3-1.amzn2023
Zookeeper	Core	hadoop-yarn-nodemanager	추가 항목	3.4.0.amzn.3-1.amzn2023
Zookeeper	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Zookeeper	Core	hadoop-yarn-resourcemanager	추가 항목	3.4.0.amzn.3-1.amzn2023
Zookeeper	Core	hadoop-yarn-timelineserver	추가 항목	3.4.0.amzn.3-1.amzn2023
Zookeeper	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

7.7.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.7.0에 대한 정보가 포함되어 있습니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.7.0 application upgrades include Delta 3.2.1-amzn-2, Hudi 0.15.0-amzn-4, Iceberg 1.7.1-amzn-0, and Trino 467.

알려진 문제

- 이 릴리스에는 적용되지 않습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태ExecutorState라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정합니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- EMR 7.7.0부터 사용 중인 Iceberg 버전은 더 이상 Java 8을 지원하지 않습니다. 또한 Iceberg는 `emr-7.7.0-java8-latest` 및 Java 8 이미지에서 제외됩니다`emr-7.7.0-spark-rapids-java8-latest`.
- Amazon EMR on EKS를 사용한 FGAC용 Lake Formation - Amazon EMR 릴리스 7.7 이상에서는 AWS Lake Formation을 활용하여 Amazon S3에서 지원하는 데이터 카탈로그 테이블에 세분화된 액세스 제어를 적용할 수 있습니다. 이 기능을 사용하면 Amazon EMR on EKS Spark 작업 내에서 읽기 쿼리에 대한 테이블, 행, 열 및 셀 수준 액세스 제어를 구성할 수 있습니다. 자세한 내용은 [세분화된 액세스 제어를 위해 AWS Lake Formation에서 Amazon EMR on EKS 사용을] 참조하세요.

-

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
609.0일			아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1일	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

7.7.0 기본 Java 버전

Amazon EMR 릴리스 7.7 이상은 Apache Livy를 제외하고 Corretto 17(OpenJDK)을 지원하는 애플리케이션에 기본적으로 Amazon Corretto 17(OpenJDK 기반)과 함께 제공됩니다.

다음 표에는 Amazon EMR 7.7.0의 애플리케이션에 대한 기본 Java 버전이 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

7.7.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.2.1-amzn-2	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.16.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.20.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.18.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.37.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.16.0	EMR S3Select 커넥터
emr-wal-cli	1.3.2	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.68.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.20.0-amzn-1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.20.0-amzn-1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.4.0-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.4.0-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.4.0-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.4.0-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.4.0-amzn-3	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.4.0-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.4.0-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.4.0-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.4.0-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.4.0-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.4.0-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.6.1-amzn-1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.6.1-amzn-1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.6.1-amzn-1	HBase 명령줄 클라이언트
hbase-rest-server	2.6.1-amzn-1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.6.1-amzn-1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.6.1-amzn-1	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-16	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-16	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-16	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-16	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-16	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-16	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-16	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.15.0-amzn-4	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.15.0-amzn-4	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.15.0-amzn-4	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.15.0-amzn-4	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.7.1-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	12.3.0	Nvidia 드라이버와 Cuda 도구키트

구성 요소	버전	설명
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.2.1	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.287-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.287-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.287-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	467-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	467-amzn-0	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
trino-client	467-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.3-amzn-1	Spark 명령줄 클라이언트.
spark-history-server	3.5.3-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.3-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.3-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	24.10.1-amzn-1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
tensorflow	2.16.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-14	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-14	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
zeppelin-server	0.11.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.2-amzn-1	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.2-amzn-1	ZooKeeper 명령줄 클라이언트.

7.7.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-7.7.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j2	Phoenix의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.

분류	설명	재구성 작업
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.

분류	설명	재구성 작업
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatchAgent service.

7.7.0 변경 로그

7.7.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2025-02-07	문서 게시	Amazon EMR 7.7.0 릴리스 정보 최초 게시
2025-02-06	배포 완료	지원되는 모든 리전에 Amazon EMR 7.7.0 완전 배포
2025-01-30	초기 제한 릴리스	Amazon EMR 7.7.0이 초기 상용 리전에 처음 배포됨

Amazon EMR 릴리스 7.6.0

7.6.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.6.0에 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜	
초기 릴리스 날짜	2025년 1월 10일	

지원 단계	날짜
표준 지원 종료 시점	2027년 1월 9일
지원 종료	2027년 1월 10일
수명 종료	2028년 1월 10일

7.6.0 릴리스 구성 요소

Amazon EMR 7.6.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta, Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, Oozie, Pig, Presto, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.6.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, Presto, Pig, Spark, TensorFlow, Tez, Trino및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
AWS Java용 SDK	2.29.25, 1.12.779	2.28.8, 1.12.772	2.28.8, 1.12.772	2.25.70, 1.12.747
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.18
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.1-amzn-1	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0
Flink	1.20.0	1.19.1	1.19.1	1.18.1
Ganglia	-	-	-	-
HBase	2.6.1	2.5.10	2.5.5	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.4.0	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-3	0.15.0-amzn-2	0.15.0-amzn-1	0.15.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-2	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	-

	emr-7.6.0	emr-7.5.0	emr-7.4.0	emr-7.3.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.2.0	5.2.0	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.287	0.287	0.287	0.285
Spark	3.5.3	3.5.2	3.5.2	3.5.1
Sqoop	-	-	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.16.1
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	457	446	446	442
Zeppelin	0.11.1	0.11.1	0.11.1	0.11.1
ZooKeeper	3.9.2	3.9.2	3.9.2	3.9.1

7.6.0 일반적인 취약성 및 노출

다음 표에는 권장 Amazon EMR 7.6.0 구성에서 실행되는 EMR 클러스터에 영향을 주지 않는 모든 CVEs가 나열되어 있습니다. Amazon EMR은 업스트림 오픈 소스의 수정 사항 가용성에 종속되어 있으며, Amazon EMR이 수정 사항을 확인한 후 90일 이내에 코어 엔진 구성 요소의 일부로 안정적인 최신 버전을 제공합니다.

이 표는 2025년 1월 10일에 업데이트되었습니다.

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-15250	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-15250

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-17521	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-17521
CVE-2024-12798	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-12798
CVE-2024-12801	낮음	https://nvd.nist.gov/vuln/detail/CVE-2024-12801
CVE-2024-47554	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-47554
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-38749	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38749
CVE-2022-38750	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38750
CVE-2022-38751	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38751
CVE-2022-38752	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38752
CVE-2022-41854	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-41854
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481
CVE-2024-36114	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-36114
CVE-2015-1832	심각	https://nvd.nist.gov/vuln/detail/CVE-2015-1832
CVE-2018-1313	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-1313
CVE-2019-10219	중간	https://nvd.nist.gov/vuln/detail/CVE-2019-10219
CVE-2020-10693	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-10693
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2020-13955	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-13955
CVE-2020-29582	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-29582
CVE-2020-8908	낮음	https://nvd.nist.gov/vuln/detail/CVE-2020-8908
CVE-2021-28170	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-28170
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-38153	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-38153
CVE-2022-24329	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-24329
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-36364	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-36364
CVE-2022-39135	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-39135
CVE-2022-42889	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-42889
CVE-2022-46337	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2022-46751	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-20863	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-20863
CVE-2023-20883	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-20883
CVE-2023-25194	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-33201	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-33201
CVE-2023-33202	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-33202
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-42503	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-42503
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-5072
CVE-2023-51074	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-51074
CVE-2024-23945	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-23945
CVE-2024-25710	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-25710

CVE ID	심각도	CVE 세부 정보 URL
CVE-2024-26308	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-26308
CVE-2024-29857	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-29857
CVE-2024-30171	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-30171
CVE-2024-31141	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-31141
CVE-2024-38808	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-38808
CVE-2024-38820	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-38820
CVE-2024-47561	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-47561
CVE-2024-52046	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-52046
CVE-2024-53990	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-53990
CVE-2024-56128	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-56128
CVE-2024-7254	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-7254
GHSA-chfm-68vv-pvw5	낮음	링크를 사용할 수 없습니다.
CVE-2024-47535	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-47535

CVE ID	심각도	CVE 세부 정보 URL
CVE-2012-5783	중간	https://nvd.nist.gov/vuln/detail/CVE-2012-5783
CVE-2012-6153	중간	https://nvd.nist.gov/vuln/detail/CVE-2012-6153
CVE-2016-5018	심각	https://nvd.nist.gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-6796
CVE-2018-10237	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-10237
CVE-2018-11798	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-11798
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2020-15522	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-15522
CVE-2020-1945	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-1945
CVE-2021-36373	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-36373
CVE-2021-36374	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-36374
CVE-2018-1000180	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-1000180
CVE-2020-26939	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-26939

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-28052	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-28052
CVE-2021-29425	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-29425
CVE-2015-5237	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-17485
CVE-2017-7525	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14721

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-19360	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086
CVE-2019-12384	중간	https://nvd.nist.gov/vuln/detail/CVE-2019-12384
CVE-2019-12814	중간	https://nvd.nist.gov/vuln/detail/CVE-2019-12814
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14893

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-16335	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11112

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-11113	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36180	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9546

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-9547	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2021-22569	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-22569
CVE-2021-22570	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-22570
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2023-50387	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-50387
CVE-2023-50868	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2023-50868
CVE-2023-52428	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-52428
CVE-2024-23454	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-23454
CVE-2024-23944	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-23944

CVE ID	심각도	CVE 세부 정보 URL
CVE-2024-25638	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-25638
CVE-2024-29025	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-29025
CVE-2024-29131	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-29131
CVE-2024-29133	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-29133
CVE-2024-51504	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-51504
CVE-2024-6763	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-6763
CVE-2024-8184	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-8184
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2013-7285	심각	https://nvd.nist.gov/vuln/detail/CVE-2013-7285

CVE ID	심각도	CVE 세부 정보 URL
CVE-2014-8125	높음	https://nvd.nist.gov/vuln/detail/CVE-2014-8125
CVE-2016-3674	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-3674
CVE-2017-7957	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-7957
CVE-2019-10173	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-10173
CVE-2020-26217	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-26217
CVE-2020-26258	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-26258
CVE-2021-21341	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-21341
CVE-2021-21342	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-21342
CVE-2021-21343	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-21343
CVE-2021-21344	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-21344
CVE-2021-21345	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-21345
CVE-2021-21346	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-21346
CVE-2021-21347	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-21347

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-21348	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-21348
CVE-2021-21349	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-21349
CVE-2021-21350	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-21350
CVE-2021-21351	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-21351
CVE-2021-29505	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-29505
CVE-2021-39139	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39139
CVE-2021-39141	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39141
CVE-2021-39144	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39144
CVE-2021-39145	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39145
CVE-2021-39146	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39146
CVE-2021-39147	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39147
CVE-2021-39148	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39148
CVE-2021-39149	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39149

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-39150	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39150
CVE-2021-39151	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39151
CVE-2021-39152	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39152
CVE-2021-39153	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39153
CVE-2021-39154	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-39154
CVE-2021-41411	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-41411
CVE-2021-43859	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-43859
CVE-2022-1415	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-1415
CVE-2022-40151	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40151
CVE-2022-41966	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-41966
CVE-2024-47072	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-47072
CVE-2019-10172	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-10172
CVE-2019-10202	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-10202

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-45688	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45688
CVE-2023-0833	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-0833
CVE-2024-21634	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-21634
CVE-2024-9823	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-9823
CVE-2022-37866	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-37866
CVE-2012-0881	높음	https://nvd.nist.gov/vuln/detail/CVE-2012-0881
CVE-2022-23437	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-23437
CVE-2020-13936	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13936
CVE-2022-37865	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-37865
CVE-2009-2625	중간	https://nvd.nist.gov/vuln/detail/CVE-2009-2625
CVE-2017-10355	중간	https://nvd.nist.gov/vuln/detail/CVE-2017-10355
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2020-14338	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-14338

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-2799	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-2799
CVE-2021-27568	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-27568
CVE-2013-4002	높음	https://nvd.nist.gov/vuln/detail/CVE-2013-4002
CVE-2020-13956	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-13956
CVE-2020-25649	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2023-26049	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-26049
CVE-2024-43382	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2024-43382
CVE-2023-34462	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-34462
CVE-2021-21290	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-21290
CVE-2023-44487	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-44487
CVE-2019-20445	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20445
GHSA-xpw8-rcwv-8f8p	높음	링크를 사용할 수 없습니다.
CVE-2021-37136	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37136

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-21295	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-21295
GHSA-58qw-p7qm-5rvh	낮음	링크를 사용할 수 없습니다.
CVE-2021-37533	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-37533
CVE-2019-20444	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20444
CVE-2022-40152	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40152

7.6.0 릴리스 구성 요소 세부 정보

Amazon EMR 7.6.0의 추가 패키지에 대한 자세한 내용은 다음 표를 참조하세요.

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	추가 항목	3.4.0.amzn.2-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	추가 항목	2.36.0-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper	추가 항목	3.9.2.amzn.0-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	추가 항목	3.9.2.amzn.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hadoop-client	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-httpfs	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-kms	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.2-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hadoop-yarn-resourcemanager	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Flink	Core	hudi	추가 항목	0.15.0.amzn.3-1.amzn2023
Flink	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Flink	Core	hadoop	추가 항목	3.4.0.amzn.2-1.amzn2023
HBase	Core	s3-dist-cp	추가 항목	2.36.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
HCatalog	Core	mariadb-connector-java	추가 항목	
Hive	Core	hadoop-httpfs	추가 항목	3.4.0.amzn.2-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Hive	Core	hive-hbase	추가 항목	3.1.3.amzn.15-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Hive	Core	mariadb-connector-java	추가 항목	2.7.2-1
Hive	Core	s3-dist-cp	추가 항목	2.36.0-1.amzn2023
피닉스	Core	s3-dist-cp	추가 항목	2.36.0-1.amzn2023
Spark	Core	spark-datanucleus	추가 항목	3.5.1.amzn.0-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.2-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	bigtop-utils	추가 항목	1.2.0-1.amzn2023
Trino	Core	hadoop-client	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-hdfs	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	추가 항목	3.4.0.amzn.2-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-hdfs-namenode	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-kms	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hive	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	추가 항목	3.1.3.amzn.12-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	mariadb-connector-java	추가 항목	2.7.2-1
Trino	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Trino	Core	zookeeper	추가 항목	3.9.2.amzn.0-1.amzn2023
Trino	Core	zookeeper-server	추가 항목	3.9.2.amzn.0-1.amzn2023
Trino	Core	bigtop-groovy	추가 항목	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	추가 항목	1.2.4-1.amzn2023
Trino	Core	hadoop	추가 항목	3.4.0.amzn.2-1.amzn2023
Trino	Core	hadoop-lzo	추가 항목	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	추가 항목	3.1.3.amzn.15-1.amzn2023
Trino	Core	hive-jdbc	추가 항목	3.1.3.amzn.12-1.amzn2023
Zookeeper	Core	emrfs	추가 항목	3.9.2.amzn.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-client	추가 항목	3.9.2.amzn.0-1.amzn2023
Zookeeper	Core	hadoop-hdfs	추가 항목	3.9.2.amzn.0-1.amzn2023
Zookeeper	Core	hadoop-hdfs-datanode	추가 항목	3.9.2.amzn.0-1.amzn2023
Zookeeper	Core	hadoop-hdfs-journalnode	추가 항목	3.9.2.amzn.0-1.amzn2023
Zookeeper	Core	hadoop-hdfs-namenode	추가 항목	3.9.2.amzn.0-1.amzn2023
Zookeeper	Core	hadoop-httpfs	추가 항목	3.9.2.amzn.0-1.amzn2023
Zookeeper	Core	hadoop-kms	추가 항목	3.9.2.amzn.0-1.amzn2023
Zookeeper	Core	hadoop-mapreduce	추가 항목	3.4.0.amzn.2-1.amzn2023
Zookeeper	Core	hadoop-mapreduce-historyserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Zookeeper	Core	hadoop-yarn	추가 항목	3.4.0.amzn.2-1.amzn2023
Zookeeper	Core	hadoop-yarn-nodemanager	추가 항목	3.4.0.amzn.2-1.amzn2023
Zookeeper	Core	hadoop-yarn-proxyserver	추가 항목	3.4.0.amzn.2-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-yarn-resourcemanager	추가 항목	3.4.0.amzn.2-1.amzn2023
Zookeeper	Core	hadoop-yarn-timelineserver	추가 항목	3.4.0.amzn.2-1.amzn2023
Zookeeper	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

7.6.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.6.0에 대한 정보가 포함되어 있습니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.6.0 application upgrades include HBase 2.6.1, Flink 1.20.0, Spark 3.5.3, and Trino 457.

알려진 문제

- 관리형 조정이 활성화되면 YARN ResourceManager(RM)에 심각한 교착 문제가 발생하여 동시 작업으로 DECOMMISSIONING에서 DECOMMISSIONED 상태로 전환할 때 응답하지 않습니다. 이는 특히 클러스터가 큰 환경에서 발생합니다. 이 경우 노드 상태 전환과 애플리케이션 상태 전환이 모두 처리되지 않아 ResourceManager의 핵심 기능을 효과적으로 마비시킵니다. 해결 방법으로 프라이머리 노드에서 다음 명령을 사용하여 RM JVM을 다시 시작합니다.

  ```
  sudo systemctl restart hadoop-yarn-resourcemanager
  ```

- `CapacityScheduler` 또는 이외의 YARN 스케줄러`FairScheduler`(예: FIFO 스케줄러)로 EMR 클러스터를 시작하면 `YARNMemoryAvailablePercentage` 지표가 CloudWatch로 푸시되지 않습니다. 이 문제는 관리형 조정이 활성화된 클러스터의 축소에 영향을 미칩니다. `YARNMemoryAvailablePercentage` 지표에 따라 달라지는 경우 사용자 지정 관리형 조정에도 영향을 미칠 수 있습니다. 해결 방법으로 `CapacityScheduler` 또는를 사용합니다`FairScheduler`.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- CloudWatch 에이전트는 EMR 7.6 이상용 런타임 역할 클러스터에서 지원됩니다. 자세한 내용은 Amazon CloudWatch 에이전트를 참조하세요.
- Amazon Elastic Map Reduce(EMR) 릴리스 버전 7.6.0부터 S3A 파일 시스템 커넥터는 이제 Amazon S3 클라이언트 측 암호화를 지원합니다.
- EMR 7.6.0부터 TLS는 HMaster 및 RegionServer 엔드포인트에서 지원됩니다. 자세한 내용은 HBase: HBase를 사용한 TLS 지원에 대한 다음 표를 참조하세요.
-

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023 년 7월 2025 년 609.0 일	6.1.140-154.222.am zn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1일	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 250123.4	6.1.124-134.200.amzn2023	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250110	6.1.119-129.201.amzn2023	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023.5.20240916	6.1.109-118.189.amzn2023	2024년 9월 6일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

7.6.0 기본 Java 버전

Amazon EMR 릴리스 7.6 이상은 Apache Livy를 제외하고 Corretto 17(OpenJDK)을 지원하는 애플리케이션에 기본적으로 Amazon Corretto 17(OpenJDK 기반)과 함께 제공됩니다.

다음 표에는 Amazon EMR 7.6.0의 애플리케이션에 대한 기본 Java 버전이 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	17, 11, 8
Zookeeper	17, 11, 8

7.6.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.2.1-amzn-1	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.15.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.19.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.18.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.36.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.15.0	EMR S3Select 커넥터
emr-wal-cli	1.3.1	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.67.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.20.0-amzn-0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.20.0-amzn-0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.4.0-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.4.0-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.4.0-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.4.0-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.4.0-amzn-2	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.4.0-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.4.0-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.4.0-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.4.0-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.4.0-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.4.0-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.6.1-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.6.1-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.6.1-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.6.1-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.6.1-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.6.1-amzn-0	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-15	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-15	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-15	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-15	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-15	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-15	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-15	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.15.0-amzn-3	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.15.0-amzn-3	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.15.0-amzn-3	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.15.0-amzn-3	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.6.1-amzn-2	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	12.3.0	Nvidia 드라이버와 Cuda 도구 키트

구성 요소	버전	설명
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.287-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.287-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.287-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	457-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	457-amzn-0	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
trino-client	457-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.3-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.5.3-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.3-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.3-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	24.10.1-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
tensorflow	2.16.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-13	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-13	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
zeppelin-server	0.11.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.2-amzn-0	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.2-amzn-0	ZooKeeper 명령줄 클라이언트.

7.6.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-7.6.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j2	Phoenix의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.

분류	설명	재구성 작업
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.

분류	설명	재구성 작업
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatc hAgent service.

7.6.0 변경 로그

7.6.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2025-01-13	문서 게시	Amazon EMR 7.6.0 릴리스 정보가 처음 게시됨
2025-01-10	배포 완료	지원되는 모든 리전에 Amazon EMR 7.6.0 완전 배포
2025-12-26	초기 릴리스	Amazon EMR 7.6.0이 초기 상용 리전에 처음 배포됨

Amazon EMR 릴리스 7.5.0

7.5.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.5.0에 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜	
초기 릴리스 날짜	2024년 11월 21일	

지원 단계	날짜	
표준 지원 종료 시점	2026년 11월 20일	
지원 종료	2026년 11월 21일	
수명 종료	2027년 11월 21일	

권장 소프트웨어 및 구성에 따라 EMR 클러스터에 영향을 주지 않는 중요하고 일반적인 취약성 및 노출(CVEs) 목록은 7.5.0 코어 엔진에 대해 알려진 일반적인 취약성 및 노출을 참조하세요.

7.5.0 릴리스 구성 요소

Amazon EMR 7.5.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta, Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, Oozie, Pig, Presto, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.5.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, Phoenix, , Pig, Spark, PrestoTrino, TensorFlowTez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
AWS Java용 SDK	2.28.8, 1.12.772	2.28.8, 1.12.772	2.25.70, 1.12.747	2.23.18, 1.12.705
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.18	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.1-amzn-0	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0
Flink	1.19.1	1.19.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.10	2.5.5	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.4.0	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-2	0.15.0-amzn-1	0.15.0-amzn-0	0.14.1-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-1	1.6.1-amzn-0	1.5.2-amzn-0	1.5.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0

	emr-7.5.0	emr-7.4.0	emr-7.3.0	emr-7.2.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	-	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.2.0	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.287	0.287	0.285	0.285
Spark	3.5.2	3.5.2	3.5.1	3.5.1
Sqoop	-	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.16.1	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	446	446	442	436
Zeppelin	0.11.1	0.11.1	0.11.1	0.10.1
ZooKeeper	3.9.2	3.9.2	3.9.1	3.9.1

7.5.0 일반적인 취약성 및 노출

다음 표에는 권장 Amazon EMR 7.5.0 구성에서 실행되는 EMR 클러스터에 영향을 주지 않는 모든 CVEs가 나열되어 있습니다. Amazon EMR은 업스트림 오픈 소스의 수정 사항 가용성에 종속되어 있으며, Amazon EMR이 수정 사항을 확인한 후 90일 이내에 코어 엔진 구성 요소의 일부로 안정적인 최신 버전을 제공합니다.

이 표는 2024년 11월 21일에 업데이트되었습니다.

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2021-37533	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37533
CVE-2018-14721	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-14721
CVE-2020-26939	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-26939
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2024-47554	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-47554
CVE-2021-38153	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-38153
CVE-2020-36188	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-10673	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-10673
CVE-2020-8908	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-8908
CVE-2022-3509	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-41966	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-41966

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-3635	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2017-17485	중간	https://nvd.nist.gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-25710
CVE-2023-25194	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-25194
GHSA-58qw-p7qm-5rvh	분류되지 않음	https://nvd.nist.gov/vuln/detail/GHSA-58qw-p7qm-5rvh
CVE-2023-51074	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-51074
CVE-2023-20883	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-20883
CVE-2022-42004	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2024-47561	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-47561
CVE-2023-2976	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2023-33202	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-33202

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-14892	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14892
CVE-2023-33201	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-33201
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-10237
CVE-2015-5237	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-5237
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2024-36114	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-36114
CVE-2024-7254	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061
CVE-2024-23944	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-23944
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2023-52428	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	낮음	https://nvd.nist.gov/vuln/detail/CVE-2024-29857

CVE ID	심각도	CVE 세부 정보 URL
CVE-2024-30171	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-30171
CVE-2022-41854	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-41854
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2018-7489	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-7489
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111
CVE-2020-15250	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-15250
CVE-2020-9548	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-51504
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2024-29025	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-29025
CVE-2024-36124	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-36124

CVE ID	심각도	CVE 세부 정보 URL
CVE-2013-4002	심각	https://nvd.nist.gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2020-36186	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2022-45688	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-45688
CVE-2020-17521	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-17521
CVE-2024-8184	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-14195
CVE-2023-40167	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-6763
CVE-2022-25647	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2018-14719	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-14719
CVE-2024-29131	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-29131
CVE-2020-9546	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-9546

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-1000180	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2021-22570	낮음	https://nvd.nist.gov/vuln/detail/CVE-2021-22570
CVE-2022-46751	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-46751
CVE-2021-36373	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-36373
CVE-2023-26048	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-30172
CVE-2022-36364	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-36364
CVE-2022-46337	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2019-16869	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-16869
CVE-2022-38751	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-38751
CVE-2022-1415	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-1415
CVE-2021-22569	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-22569

CVE ID	심각도	CVE 세부 정보 URL
CVE-2024-35255	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-35255
CVE-2021-28170	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-28170
CVE-2022-40149	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2023-50868	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-50868
CVE-2023-34454	high	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2022-38750	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38750
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-42503
CVE-2024-25638	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-25638
CVE-2024-29133	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-29133
CVE-2022-37865	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-37865
CVE-2018-19361	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-19362

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-10650	보통	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	보통	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2024-23454	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2020-24750	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-11112
CVE-2019-14439	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2023-34455	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2024-26308	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-20445
CVE-2023-0833	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-38752	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38752
CVE-2020-1945	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-1945
CVE-2020-13936	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-36478
CVE-2020-10968	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-10968
CVE-2020-28052	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-43859
CVE-2021-27568	낮음	https://nvd.nist.gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	낮음	https://nvd.nist.gov/vuln/detail/CVE-2020-11113
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2023-5072	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2018-11307
CVE-2024-21634	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-21634

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2020-36180	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538
CVE-2021-29425	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-29425
CVE-2020-8840	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2019-16943	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2016-5018	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-34447
CVE-2024-39676	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-39676
CVE-2021-21290	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-14718	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14718
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181
CVE-2022-39135	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2022-39135
CVE-2019-16335	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2019-16335
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2020-14062	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-15522
CVE-2019-14540	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2017-15095	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14893
CVE-2019-12814	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12814

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-10672	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-24616	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-13949	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-50387
CVE-2018-12022	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-12022
CVE-2020-36189	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-12384
CVE-2023-26049	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2023-26049
CVE-2022-1471	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2022-40152	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40152

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-20863	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-11798
CVE-2019-20330	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-17531
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2020-11619	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11619
CVE-2022-40151	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40151
CVE-2019-17267	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2018-10054	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-10054

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-14720	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-14720
CVE-2022-24329	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2022-24329
CVE-2020-36183	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2022-37866	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-37866
CVE-2015-1832	심각	https://nvd.nist.gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2018-1313
CVE-2018-12023	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12023
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2024-38808	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-20444
CVE-2012-0881	높음	https://nvd.nist.gov/vuln/detail/CVE-2012-0881

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-13956	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13956
CVE-2022-38749	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-38749
CVE-2017-7525	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2017-7525
CVE-2018-19360	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2018-19360
CVE-2021-37136	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2021-37136
CVE-2022-42889	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42889
CVE-2021-36374	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13955
CVE-2022-25857	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-25857

7.5.0 릴리스 구성 요소 세부 정보

Amazon EMR 7.5.0의 추가 패키지에 대한 자세한 내용은 다음 표를 참조하세요.

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hudi	추가 항목	0.15.0.amzn.0-1.amzn2023
Flink	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Flink	Core	hadoop	추가 항목	3.3.6.amzn.5-1.amzn2023
HBase	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
HCatalog	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
HCatalog	Core	mariadb-connector-java	추가 항목	
Hive	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Hive	Core	hive-hbase	추가 항목	3.1.3.amzn.12-1.amzn2023
Hive	Core	mariadb-connector-java	추가 항목	2.7.2-1
Hive	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
피닉스	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
Spark	Core	spark-datanucleus	추가 항목	3.5.1.amzn.0-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Tez의 Hive 실행 시간 비교	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	bigtop-utils	추가 항목	1.2.0-1.amzn2023
Trino	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hive	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	mariadb-connector-java	추가 항목	2.7.2-1
Trino	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Trino	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	추가 항목	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	추가 항목	1.2.4-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-lzo	추가 항목	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-jdbc	추가 항목	3.1.3.amzn.12-1.amzn2023
Zookeeper	Core	emrfs	추가 항목	2.64.0-1.amzn2023
Zookeeper	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

7.5.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.5.0에 대한 정보가 포함되어 있습니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.5.0 application upgrades include Delta 3.2.1-amzn-0, HBase 2.5.10, Hudi 0.15.0-amzn-2, and Iceberg 1.6.1-amzn-1.
- EMR 미리 쓰기 로그(WAL)는 클러스터 간 복제를 지원합니다. 자세한 내용은 EMR WAL 클러스터 간 복제를 참조하세요.

알려진 문제

- Amazon EMR 릴리스 7.5 이상 릴리스에서 Apache Sqoop이 제거되었습니다.
- 관리형 조정이 활성화되면 YARN ResourceManager(RM)에 심각한 교착 문제가 발생하여 동시 작업으로 DECOMMISSIONING에서 DECOMMISSIONED 상태로 전환할 때 응답하지 않습니다. 이는 특히 클러스터가 큰 환경에서 발생합니다. 이 경우 노드 상태 전환과 애플리케이션 상태 전환이 모두 처리되지 않아 ResourceManager의 핵심 기능을 효과적으로 마비시킵니다. 해결 방법으로 프라이머리 노드에서 다음 명령을 사용하여 RM JVM을 다시 시작합니다.

```
sudo systemctl restart hadoop-yarn-resourcemanager
```

- `CapacityScheduler` 또는 이외의 YARN 스케줄러`FairScheduler`(예: FIFO 스케줄러)로 EMR 클러스터를 시작하면 `YARNMemoryAvailablePercentage` 지표가 CloudWatch로 푸시되지 않습니다. 이 문제는 관리형 조정이 활성화된 클러스터의 축소에 영향을 미칩니다. `YARNMemoryAvailablePercentage` 지표에 따라 달라지는 경우 사용자 지정 관리형 조정에도 영향을 미칠 수 있습니다. 해결 방법으로 `CapacityScheduler` 또는를 사용합니다`FairScheduler`.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 TRUE (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.

- Amazon EMR은 클러스터에 대해 선택한 인스턴스 유형에 따라 TensorFlow 라이브러리의 다양한 빌드를 사용합니다. 또한 Amazon EMR은 EMR-7.5.0 이상의 aarch64 인스턴스 유형(Graviton 인스턴스)이 있는 클러스터에 TensorFlow를 지원합니다.

  다음 표에는 EMR-7.5.0의 인스턴스 유형별 빌드가 나열되어 있습니다.

EC2 인스턴스 유형	TensorFlow 빌드
M5 및 C5	Intel MKL 최적화가 적용된 Tensorflow 2.16.1
P2, P4D, P5, G4DN, G5, G6, GR6	CUDA 12.3, cuDNN 8.9.7.29가 적용된 Tensorflow 2.16.1
P3, P3DN, G3, G3S	CUDA 12.3, cuDNN 8.9.7.29, NCCL 2.20.3-1 Nvidia NCCL이 포함된 Tensorflow 2.16.1은 P3 인스턴스에서만 사용할 수 있습니다. 최종 사용자 라이선스 계약(EULA) - Amazon EMR에서 Nvidia 구성 요소를 사용하면 제품 EULA에 설명된 이용 약관에 동의하는 것입니다.
Graviton(EMR-7.5.0 이상)을 포함한 기타 모든 항목	Tensorflow 2.16.1

- FGAC(Fine-Grained Access Control)가 활성화된 AWS Lake Formation Amazon EMR 클러스터는 이제 다음 hdfs-site.xml 구성으로 최대 256개의 동시 단계를 지원할 수 있습니다.

```
[
  {
    "Classification": "hdfs-site",
    "Properties": {
      "dfs.datanode.handler.count": "40",
      "dfs.datanode.max.transfer.threads": "16384",
      "dfs.namenode.handler.count": "64",
      "dfs.replication": "1"
    }
  },
  {
    "Classification": "hdfs-env",
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "5120"
        }
      }
    ],
    "Properties": {}
  }
]
```

-

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1일	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 250123.4	6.1.124-134.200.amzn2023	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250110	6.1.119-129.201.amzn2023	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023.5.20240916	6.1.109-118.189.amzn2023	2024년 9월 6일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

7.5.0 기본 Java 버전

Amazon EMR 릴리스 7.5 이상은 Apache Livy를 제외하고 Corretto 17(OpenJDK)을 지원하는 애플리케이션에 기본적으로 Amazon Corretto 17(OpenJDK 기반)과 함께 제공됩니다.

다음 표에는 Amazon EMR 7.5.0의 애플리케이션에 대한 기본 Java 버전이 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	8
Zookeeper	17, 11, 8

7.5.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.2.1-amzn-0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.14.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.18.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.16.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.35.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.14.0	EMR S3Select 커넥터
emr-wal-cli	1.3.1	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.66.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.19.1-amzn-1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.19.1-amzn-1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.4.0-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.4.0-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.4.0-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.4.0-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.4.0-amzn-1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.4.0-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.4.0-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.4.0-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.4.0-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.4.0-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.4.0-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.5.10-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.5.10-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.5.10-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.5.10-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.5.10-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.5.10-amzn-0	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-14	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-14	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-14	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-14	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-14	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-14	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-14	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.15.0-amzn-2	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.15.0-amzn-2	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.15.0-amzn-2	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.15.0-amzn-2	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.6.1-amzn-1	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	12.3.0	Nvidia 드라이버와 Cuda 도구키트

구성 요소	버전	설명
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.287-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.287-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.287-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	446-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	446-amzn-0	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
trino-client	446-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.2-amzn-1	Spark 명령줄 클라이언트.
spark-history-server	3.5.2-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.2-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.2-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	24.08.1-amzn-1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
tensorflow	2.16.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-12	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-12	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
zeppelin-server	0.11.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.2-amzn-0	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.2-amzn-0	ZooKeeper 명령줄 클라이언트.

7.5.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-7.5.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j2	Phoenix의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.

분류	설명	재구성 작업
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.

분류	설명	재구성 작업
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatc hAgent service.

7.5.0 변경 로그

7.5.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2024-11-21	문서 게시	Amazon EMR 7.5.0 릴리스 정보 최초 게시
2024-11-21	배포 완료	지원되는 모든 리전에 Amazon EMR 7.5.0 완전 배포
2024-11-13	초기 릴리스	Amazon EMR 7.5.0이 초기 상용 리전에 처음 배포됨

Amazon EMR 릴리스 7.4.0

7.4.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.4.0에서 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜	
초기 릴리스 날짜	2024년 11월 13일	

지원 단계	날짜	
표준 지원 종료 시점	2026년 11월 12일	
지원 종료	2026년 11월 13일	
수명 종료	2027년 11월 13일	

권장 소프트웨어 및 구성에 따라 EMR 클러스터에 영향을 미치지 않는 중요하고 일반적인 취약성 및 노출(CVEs) 목록은 코어 엔진에 대해 알려진 7.4.0 일반적인 취약성 및 노출을 참조하세요.

7.4.0 릴리스 구성 요소

Amazon EMR 7.4.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta, Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, Oozie, Pig, Presto, Sqoop, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.4.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hive, Hadoop, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, , Phoenix, Pig, Presto, Sqoop, SparkTrino, TensorFlowTez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
AWS Java용 SDK	2.28.8, 1.12.772	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9, 3.11
Scala	2.12.18	2.12.18	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300032.2
Delta	3.2.0-amzn-1	3.2.0-amzn-0	3.1.0	3.0.0
Flink	1.19.1	1.18.1	1.18.1	1.18.1
Ganglia	-	-	-	-
HBase	2.5.5	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.4.0	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-1	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.6.1-amzn-0	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0

	emr-7.4.0	emr-7.3.0	emr-7.2.0	emr-7.1.0
Livy	0.8.0	0.8.0	0.8.0	0.8.0
MXNet	-	-	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.2.0	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.287	0.285	0.285	0.284
Spark	3.5.2	3.5.1	3.5.1	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.16.1	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	446	442	436	435
Zeppelin	0.11.1	0.11.1	0.10.1	0.10.1
ZooKeeper	3.9.2	3.9.1	3.9.1	3.9.1

7.4.0 일반적인 취약성 및 노출

다음 표에는 권장 Amazon EMR 7.4.0 구성에서 실행되는 EMR 클러스터에 영향을 주지 않는 모든 CVEs가 나열되어 있습니다. Amazon EMR은 업스트림 오픈 소스의 수정 사항 가용성에 종속되어 있으며, Amazon EMR이 수정 사항을 확인한 후 90일 이내에 코어 엔진 구성 요소의 일부로 안정적인 최신 버전을 제공합니다.

이 표는 2024년 11월 15일에 업데이트되었습니다.

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-37533	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37533
CVE-2018-14721	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-14721
CVE-2020-26939	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-26939
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2024-47554	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-47554
CVE-2021-38153	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-38153
CVE-2022-41966	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-41966
CVE-2023-3635	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2020-25649	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2017-17485	중간	https://nvd.nist.gov/vuln/detail/CVE-2017-17485
CVE-2024-25710	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-25710
CVE-2023-25194	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-25194

CVE ID	심각도	CVE 세부 정보 URL
GHSA-58qw-p7qm-5rvh	분류되지 않음	https://nvd.nist.gov/vuln/detail/GHSA-58qw-p7qm-5rvh
CVE-2023-51074	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-51074
CVE-2023-20883	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-20883
CVE-2022-42004	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2024-47561	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-47561
CVE-2023-2976	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2023-33202	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-33202
CVE-2019-14892	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14892
CVE-2023-33201	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-33201
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2018-10237	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-10237
CVE-2015-5237	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-5237

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2024-36114	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-36114
CVE-2024-7254	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-7254
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061
CVE-2024-23944	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-23944
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2023-52428	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-52428
CVE-2024-29857	낮음	https://nvd.nist.gov/vuln/detail/CVE-2024-29857
CVE-2024-30171	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-30171
CVE-2022-41854	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-41854
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2018-7489	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-7489
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111
CVE-2020-15250	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-15250
CVE-2020-9548	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-9548
CVE-2024-51504	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-51504
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2024-29025	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-29025
CVE-2024-36124	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-36124
CVE-2013-4002	심각	https://nvd.nist.gov/vuln/detail/CVE-2013-4002
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2020-36186	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2022-45688	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-45688
CVE-2020-17521	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-17521

CVE ID	심각도	CVE 세부 정보 URL
CVE-2024-8184	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-8184
CVE-2020-14195	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-14195
CVE-2023-40167	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-40167
CVE-2024-6763	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-6763
CVE-2022-25647	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2018-14719	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-14719
CVE-2024-29131	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-29131
CVE-2020-9546	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-9546
CVE-2018-1000180	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-1000180
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2021-22570	낮음	https://nvd.nist.gov/vuln/detail/CVE-2021-22570
CVE-2022-46751	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-46751
CVE-2021-36373	중간	https://nvd.nist.gov/vuln/detail/CVE-2021-36373

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-26048	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-26048
CVE-2024-30172	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-30172
CVE-2022-36364	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-36364
CVE-2022-46337	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2019-16869	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-16869
CVE-2022-38751	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-38751
CVE-2022-1415	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-1415
CVE-2021-22569	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-22569
CVE-2024-35255	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-35255
CVE-2021-28170	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-28170
CVE-2022-40149	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2023-50868	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-50868
CVE-2022-38750	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38750

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2023-42503	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-42503
CVE-2024-25638	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-25638
CVE-2024-29133	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-29133
CVE-2022-37865	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-37865
CVE-2018-19361	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-19362
CVE-2020-10650	보통	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-35490	보통	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2024-23454	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-23454
CVE-2023-43642	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2020-24750	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-11112	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-11112

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-14439	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2023-34455	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2024-26308	중간	https://nvd.nist.gov/vuln/detail/CVE-2024-26308
CVE-2019-20445	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-20445
CVE-2023-0833	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-0833
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481
CVE-2012-5783	높음	https://nvd.nist.gov/vuln/detail/CVE-2012-5783
CVE-2022-38752	중간	https://nvd.nist.gov/vuln/detail/CVE-2022-38752
CVE-2020-1945	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-1945
CVE-2020-13936	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-13936
CVE-2023-36478	중간	https://nvd.nist.gov/vuln/detail/CVE-2023-36478
CVE-2020-10968	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-10968

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-28052	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-28052
CVE-2021-43859	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-43859
CVE-2021-27568	낮음	https://nvd.nist.gov/vuln/detail/CVE-2021-27568
CVE-2020-11113	낮음	https://nvd.nist.gov/vuln/detail/CVE-2020-11113
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2023-5072	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-5072
CVE-2018-11307	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-11307
CVE-2019-16943	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2022-39135	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-39135
CVE-2019-16335	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16335
CVE-2017-7525	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-7525
CVE-2018-19360	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19360
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2024-21634	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-21634
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2012-6153	낮음	https://nvd.nist.gov/vuln/detail/CVE-2012-6153
CVE-2020-36180	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538
CVE-2021-29425	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-29425
CVE-2020-8840	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-9547	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2016-5018	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-5018
CVE-2024-34447	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-34447
CVE-2024-39676	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-39676

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-21290	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-21290
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086
CVE-2018-14718	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14718
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2020-14062	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-15522	중간	https://nvd.nist.gov/vuln/detail/CVE-2020-15522
CVE-2019-14540	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2017-15095	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-15095
CVE-2018-5968	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-5968
CVE-2019-14893	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14893
CVE-2019-12814	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12814

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-10672	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-24616	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-13949	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2023-50387	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-50387
CVE-2018-12022	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-12022
CVE-2020-36189	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2021-37137	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-37137
CVE-2019-12384	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-12384
CVE-2023-26049	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2023-26049
CVE-2022-1471	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2019-16942	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2022-40152	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40152

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-20863	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-20863
CVE-2018-11798	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-11798
CVE-2019-20330	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2019-17531	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-17531
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2020-11619	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11619
CVE-2022-40151	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40151
CVE-2019-17267	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2016-6796	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2018-10054	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-10054

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-14720	중간	https://nvd.nist.gov/vuln/detail/CVE-2018-14720
CVE-2022-24329	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2022-24329
CVE-2020-36183	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2022-37866	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-37866
CVE-2015-1832	심각	https://nvd.nist.gov/vuln/detail/CVE-2015-1832
CVE-2023-38647	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-38647
CVE-2018-1313	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2018-1313
CVE-2018-12023	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12023
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2024-38808	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-38808
CVE-2019-20444	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-20444
GHSA-chfm-68vv-pvw5	높음	https://github.com/advisories/GHSA-chfm-68vv-pvw5

CVE ID	심각도	CVE 세부 정보 URL
CVE-2012-0881	높음	https://nvd.nist.gov/vuln/detail/CVE-2012-0881
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-13956	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13956
CVE-2022-38749	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-38749
CVE-2021-37136	분류되지 않음	https://nvd.nist.gov/vuln/detail/CVE-2021-37136
CVE-2022-42889	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42889
CVE-2021-36374	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36374
CVE-2020-13955	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13955
CVE-2022-25857	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-25857

7.4.0 릴리스 구성 요소 세부 정보

Amazon EMR 7.4.0의 추가 패키지에 대한 자세한 내용은 다음 표를 참조하세요.

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hudi	추가 항목	0.15.0.amzn.0-1.amzn2023
Flink	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Flink	Core	hadoop	추가 항목	3.3.6.amzn.5-1.amzn2023
HBase	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
HCatalog	Core	mariadb-connector-java	추가 항목	
Hive	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Hive	Core	hive-hbase	추가 항목	3.1.3.amzn.12-1.amzn2023
Hive	Core	mariadb-connector-java	추가 항목	2.7.2-1
Hive	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
피닉스	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
Spark	Core	spark-datanucleus	추가 항목	3.5.1.amzn.0-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	bigtop-utils	추가 항목	1.2.0-1.amzn2023
Trino	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hive	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	mariadb-connector-java	추가 항목	2.7.2-1
Trino	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Trino	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	추가 항목	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	추가 항목	1.2.4-1.amzn2023
Trino	Core	hadoop	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-lzo	추가 항목	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-jdbc	추가 항목	3.1.3.amzn.12-1.amzn2023
Zookeeper	Core	emrfs	추가 항목	2.64.0-1.amzn2023
Zookeeper	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

7.4.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.4.0에 대한 정보가 포함되어 있습니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.4.0 application upgrades include AWS SDK for Java 2.28.8, Delta 3.2.0-amzn-1, Flink 1.19.1, Hadoop 3.4.0, HBase 2.5.5, Hudi 0.15.0-amzn-1, Iceberg 1.6.1-amzn-0, Phoenix 5.2.0, Presto 0.287, Spark 3.5.2, Trino 446, and Zookeeper 3.9.2.

- 이 릴리스에는 네트워크를 통한 전송 중 데이터 암호화를 지원하는 7개의 오픈 소스 엔드포인트가 추가되었습니다. 이 개선 사항에는 Hue, Livy, JupyterEnterpriseGateway, Zookeeper 및 Ranger KMS와 같은 오픈 소스 프레임워크가 포함됩니다. 전송 중 데이터 암호화를 활성화하고 이러한 엔드포인트에 연결하려는 경우 TLS 통신을 지원하도록 코드를 수정해야 할 수 있습니다. 필요한 경우 이러한 엔드포인트에 대한 TLS를 비활성화할 수 있습니다. 지원되는 엔드포인트의 전체 목록은 전송 중 암호화 이해를 참조하세요.

알려진 문제

- emr-6.11.0 이후 이전에 Hive 셔플 핸들러를 지원한 Yarn ResourceManager의 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` 속성은 이제 emr-7.4.0에서 Yarn에서 Spark를 실행할 때 Spark 외부 셔플 서비스가 제공하는 파일에도 정상적인 폐기를 적용합니다.

  Spark on Yarn을 실행할 때 Spark 외부 셔플 서비스의 기본 동작은 애플리케이션이 종료될 때 Yarn NodeManager가 애플리케이션 셔플 파일을 제거하는 것입니다. 이는 노드 폐기 속도 및 컴퓨팅 사용률에 영향을 미칠 수 있습니다. 장기 실행 애플리케이션의 경우를 `spark.shuffle.service.removeShuffle` 로 설정`true`하여 더 이상 사용되지 않는 셔플 파일을 제거하면 활성 셔플 데이터 없이 노드를 더 빠르게 폐기할 수 있습니다.

- Amazon S3 Select 가용성 - Amazon S3 Select는 Presto 및 Trino에 대한 리전 간 쿼리를 지원하지 않습니다. Presto 및 Trino 릴리스에 사용할 수 있는 특정 기능에 대한 자세한 내용은 Presto 및 Trino를 참조하세요.
- Amazon EMR 릴리스 7.4에서는 Apache Ranger가 지원되지 않습니다.
- Apache Sqoop은 알려진 문제 SQOOP-3485로 인해 Apache Hadoop 3.4와 호환되지 않습니다. Amazon EMR 릴리스 7.4에는 Apache Sqoop이 포함되지만 지원되지 않습니다. Apache Sqoop은 Amazon EMR 릴리스 7.5 이상 릴리스에서 제거됩니다.
- 관리형 조정이 활성화되면 YARN ResourceManager(RM)에 심각한 교착 문제가 발생하여 동시 작업으로 DECOMMISSIONING에서 DECOMMISSIONED 상태로 전환할 때 응답하지 않습니다. 이는 특히 클러스터가 큰 환경에서 발생합니다. 이 경우 노드 상태 전환과 애플리케이션 상태 전환이 모두 처리되지 않아 ResourceManager의 핵심 기능을 효과적으로 마비시킵니다. 해결 방법으로 프라이머리 노드에서 다음 명령을 사용하여 RM JVM을 다시 시작합니다.

  ```
  sudo systemctl restart hadoop-yarn-resourcemanager
  ```

- `CapacityScheduler` 또는 이외의 YARN 스케줄러`FairScheduler`(예: FIFO 스케줄러)로 EMR 클러스터를 시작하면 `YARNMemoryAvailablePercentage` 지표가 CloudWatch로

푸시되지 않습니다. 이 문제는 관리형 조정이 활성화된 클러스터의 축소에 영향을 미칩니다. `YARNMemoryAvailablePercentage` 지표에 따라 달라지는 경우 사용자 지정 관리형 조정에도 영향을 미칠 수 있습니다. 해결 방법으로 `CapacityScheduler` 또는를 사용합니다`FairScheduler`.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Amazon EMR은 클러스터에 대해 선택한 인스턴스 유형에 따라 다른 빌드의 TensorFlow 라이브러리를 사용합니다. 또한 Amazon EMR은 EMR-7.5.0 이상의 aarch64 인스턴스 유형(Graviton 인스턴스)이 있는 클러스터에 TensorFlow를 지원합니다.
- Iceberg 워크로드에 대한 S3 요청 크기 조정 최적화 - EMR 34 v7.4 릴리스와 함께 Iceberg를 사용하는 고객의 경우 S3는 기본 엔트로피 체계를 20자 base-2(예: 35 01010110100110110010) 해시로 업데이트했습니다. 이 새로운 base-2 엔트로피 체계는 문자 범위를 줄이고 S3가 대상 Iceberg 워크로드의 수요에 맞게 요청 용량을 더 빠르게 자동으로 조정할 수 있도록 합니다. 또한 워크로드가 503 스로틀 응답을 관찰하는 시간을 줄입니다.
-

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
609.0일			타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2일	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250124	6.1.124-134.200.amzn2023	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250110	6.1.119-129.201.amzn2023	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 11월 21일	6.1.115-126.197.amzn2023	2024년 12월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023.5.20240916	6.1.109-118.189.amzn2023	2024년 9월 6일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

7.4.0 기본 Java 버전

Amazon EMR 릴리스 7.4 이상은 Apache Livy를 제외하고 Corretto 17(OpenJDK)을 지원하는 애플리케이션에 기본적으로 Amazon Corretto 17(OpenJDK 기반)과 함께 제공됩니다.

다음 표에는 Amazon EMR 7.4.0의 애플리케이션에 대한 기본 Java 버전이 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	17, 11, 8
Pig	8
Zookeeper	8

7.4.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.2.0-amzn-1	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.13.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.17.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.17.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.34.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.13.0	EMR S3Select 커넥터
emr-wal-cli	1.3.1	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.65.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.19.1-amzn-0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.19.1-amzn-0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.4.0-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.4.0-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.4.0-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.4.0-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.4.0-amzn-0	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.4.0-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.4.0-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.4.0-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.4.0-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.4.0-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.4.0-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.5.5-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.5.5-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.5.5-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.5.5-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.5.5-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.5.5-amzn-0	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-13	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-13	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-13	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-13	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-13	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-13	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-13	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.15.0-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.15.0-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.15.0-amzn-1	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.15.0-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.6.1-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	12.3.0	Nvidia 드라이버와 Cuda 도구 키트

구성 요소	버전	설명
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.287-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.287-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.287-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	446-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	446-amzn-0	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
trino-client	446-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.2-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.5.2-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.2-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.2-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	24.08.1-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.16.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-11	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
tez-on-worker	0.10.2-amzn-11	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
zeppelin-server	0.11.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.2-amzn-0	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.2-amzn-0	ZooKeeper 명령줄 클라이언트.

7.4.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 [실행 중인 클러스터에서 인스턴스 그룹 재구성](#) 단원을 참조하십시오.

emr-7.4.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.

분류	설명	재구성 작업
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j2	Phoenix의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatchAgent service.

7.4.0 변경 로그

7.4.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2024-11-18	문서 게시	Amazon EMR 7.4.0 릴리스 정보 최초 게시
2024-11-13	배포 완료	지원되는 모든 리전에 Amazon EMR 7.4.0 완전 배포
2024-10-31	배포 시작	Amazon EMR 7.4.0이 초기 상용 리전에 처음 배포됨

Amazon EMR 릴리스 7.3.0

7.3.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.3.0에 대해 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜	
초기 릴리스 날짜	2024년 10월 16일	
표준 지원 종료 시점	2026년 10월 15일	
지원 종료	2026년 10월 16일	
수명 종료	2027년 10월 16일	

권장 소프트웨어 및 구성에 따라 EMR 클러스터에 영향을 미치지 않는 중요한 상위의 일반적인 취약성 및 노출(CVE) 목록은 7.3.0 코어 엔진에 대해 알려진 일반적인 취약성 및 노출을 참조하세요.

7.3.0 릴리스 구성 요소

Amazon EMR 7.3.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta. Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, Oozie, Pig, Presto, Sqoop, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.3.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, Oozie, , Phoenix, Pig, Presto, Sqoop, SparkTensorFlowTez, Trino, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS Java용 SDK	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.285	0.285	0.284	0.283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

7.3.0 일반적인 취약성 및 노출

다음 표에는 Amazon EMR 7.3.0의 권장 구성에서 실행되는 EMR 클러스터에 영향을 주지 않는 모든 CVE가 나와 있습니다. Amazon EMR은 업스트림 오픈 소스의 수정 사항 가용성에 종속되어 있으며, Amazon EMR이 수정 사항을 확인한 후 90일 이내에 코어 엔진 구성 요소의 일부로 안정적인 최신 버전을 제공합니다.

이 표는 2024년 10월 4일에 업데이트되었습니다.

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-1000180	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-1000180
CVE-2020-28052	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-28052
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-40152	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40152

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2023-50387	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-50387

CVE ID	심각도	CVE 세부 정보 URL
CVE-2012-0881	높음	https://nvd.nist.gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	높음	https://nvd.nist.gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	심각	https://nvd.nist.gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19361

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-19362	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19362
CVE-2018-7489	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17531

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-20330	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13936
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2020-35490	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9548

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-36364	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-36364

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-37865	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-37865
CVE-2022-37866	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-37866
CVE-2022-39135	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-39135
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-46337	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2022-46751	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-34455	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2016-5018	심각	https://nvd.nist.gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2015-2156	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-18640
CVE-2017-18640	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-18640
CVE-2019-16869	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-16869

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-20444	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20445
CVE-2021-37136	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37137
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-36478
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-45688	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45688
CVE-2022-46337	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2023-5072	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-21634
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481

7.3.0 릴리스 구성 요소 세부 정보

Amazon EMR 7.3.0의 추가 항목 패키지에 대한 자세한 내용은 다음 표를 참조하세요.

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hudi	추가 항목	0.15.0.amzn.0-1.amzn2023
Flink	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Flink	Core	hadoop	추가 항목	3.3.6.amzn.5-1.amzn2023
HBase	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
HCatalog	Core	mariadb-connector-java	추가 항목	
Hive	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Hive	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Hive	Core	hive-hbase	추가 항목	3.1.3.amzn.12-1.amzn2023
Hive	Core	mariadb-connector-java	추가 항목	2.7.2-1
Hive	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
피닉스	Core	s3-dist-cp	추가 항목	2.33.0-1.amzn2023
Spark	Core	spark-datanucleus	추가 항목	3.5.1.amzn.0-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	bigtop-utils	추가 항목	1.2.0-1.amzn2023
Trino	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hive	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-hcatalog-server	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	mariadb-connector-java	추가 항목	2.7.2-1
Trino	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	추가 항목	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	추가 항목	1.2.4-1.amzn2023
Trino	Core	hadoop	추가 항목	3.3.6.amzn.5-1.amzn2023
Trino	Core	hadoop-lzo	추가 항목	0.4.19-1.amzn2023
Trino	Core	hive-hcatalog	추가 항목	3.1.3.amzn.12-1.amzn2023
Trino	Core	hive-jdbc	추가 항목	3.1.3.amzn.12-1.amzn2023
Zookeeper	Core	emrfs	추가 항목	2.64.0-1.amzn2023
Zookeeper	Core	hadoop-client	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-kms	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.5-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.5-1.amzn2023
Zookeeper	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

7.3.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.3.0에 대한 정보가 포함됩니다. 변경 사항은 7.3.0에 관련됩니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.3.0 application upgrades include Delta 3.2.0-amzn-0, Hudi 0.15.0-amzn-0, TensorFlow 2.16.1, Trino 442, and Zeppelin 0.11.1.
- 이 릴리스에는 네트워크를 통한 전송 중 데이터 암호화를 지원하는 22개의 오픈 소스 엔드포인트가 추가되었습니다. 이 개선 사항에는 Hadoop, Hive, Tez, HBase, Phoenix 및 Flink와 같은 오픈 소스 프레임워크가 포함됩니다. 전송 중 데이터 암호화를 활성화하고 이러한 엔드포인트에 연결하려는 경우 TLS 통신을 지원하도록 코드를 수정해야 할 수 있습니다. 필요한 경우 이러한 엔드포인트에 대한 TLS를 비활성화할 수 있습니다. 지원되는 엔드포인트의 전체 목록은 전송 중 암호화 이해를 참조하세요.
- Amazon EMR 7.3.0을 사용하면 Iceberg를 사용하는 Apache Hive가 이제 AWS Glue 데이터 카탈로그를 메타스토어로 지원하므로 메타데이터를 저장하고, 데이터 소스를 자동으로 크롤링하고, 메타데이터를 추출하고, 원활한 워크플로를 위해 다른 서비스와 통합하는 중앙 집중식 리포지토리와 같은 이점을 누릴 수 있습니다.

알려진 문제

- Python 3.11은 EMR Studio에서 지원되지 않습니다.
- 전송 중 암호화 및 Kerberos 인증을 사용하는 경우 Hive 메타스토어에 의존하는 Spark 애플리케이션을 사용할 수 없습니다. 이 문제는 HIVE-16340에서 수정되었으며, 오픈 소스 Spark를 Hive 3으로 업그레이드할 수 있는 경우 SPARK-44114에서 완전히 해결됩니다. 그동안 `hive.metastore.use.SSL`을 `false`로 설정하여 이 문제를 일시적으로 해결할 수 있습니다. 자세한 내용은 애플리케이션 구성을 참조하세요.

- Kerberos 인증을 사용하지 않는 경우 SSL을 사용하도록 Spark를 구성해야 합니다. 이를 수행하려면 클러스터를 프로비저닝할 때 다음 구성을 설정합니다.

```
[{
    "Classification": "spark-defaults",
    "Properties": {
        "spark.hadoop.hive.metastore.truststore.path": "/usr/share/aws/emr/
security/conf/keystore-thrift.jks",
        "spark.hadoop.hive.metastore.use.SSL": "true"
    }
}]
```

- 사용자 지정 도메인을 사용하는 경우 AWS 콘솔의 Spark 기록 서버 UI에 있는 일부 링크가 작동하지 않을 수 있습니다. 다음 스크립트를 S3 버킷에 업로드한 다음, 쉘 스크립트를 사용하도록 단계를 추가합니다.

```
#!/bin/bash

if grep -q ":8042" /etc/nginx/nginx.conf; then
    sudo sed -i '=s#\^(\\/nm\\/(?<nodeAddress>.*:8042)\\/(?<url>.*)\$)#^(\\/nm\\/
(?<nodeAddress>(\\.|-|[a-zA-Z0-9])*:8042)\\/(?<url>(node|conf|logs|stacks|jmx).*)\
$)#g' /etc/nginx/nginx.conf
else
    sudo sed -i 's#\^(\\/nm\\/(?<nodeAddress>.*:8044)\\/(?<url>.*)\$)#^(\\/nm\\/
(?<nodeAddress>(\\.|-|[a-zA-Z0-9])*:8044)\\/(?<url>(node|conf|logs|stacks|jmx).*)\
$)#g' /etc/nginx/nginx.conf
fi

sudo systemctl restart nginx
```

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 TRUE (기본값)로 설정합니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및

`spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.

- YARN에 대한 TLS 호스트 이름 확인을 지원하기 위해 이 릴리스에서는 `yarn.resourcemanager.hostname` 구성을 기본 인스턴스 IP 주소에서 기본 인스턴스 프라이빗 DNS 이름으로 변경합니다.
- MXNet 프로젝트 수명이 다하여 이제 아카이브됩니다. the section called “emr-7.2.0”는 MXNet을 포함하는 마지막 릴리스입니다.

- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1일	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250124	6.1.124-134.200.amzn2023	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250110	6.1.119-129.201.amzn2023	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 11월 21일	6.1.115-126.197.amzn2023	2024년 12월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 1031.0	6.1.112-124.190.amzn2023	2024년 11월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023.5.20240916	6.1.109-118.189.amzn2023	2024년 9월 6일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

7.3.0 기본 Java 버전

Amazon EMR 릴리스 7.3 이상에서는 Apache Livy를 제외한 Corretto 17(JDK 17)을 지원하는 애플리케이션에 대해 기본적으로 Amazon Corretto 17(OpenJDK 기반)이 함께 제공됩니다.

다음 표에서는 Amazon EMR 7.3.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	17, 11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	17, 11, 8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	8
Pig	8
Zookeeper	17, 11, 8

7.3.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.2.0-amzn-0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.12.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.16.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.16.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.33.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.12.0	EMR S3Select 커넥터
emr-wal-cli	1.3.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.64.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.18.1-amzn-2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.18.1-amzn-2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.3.6-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.6-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.6-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.6-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.6-amzn-5	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.6-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.6-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.6-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.6-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.6-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.6-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.4.17-amzn-7	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-7	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-7	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-7	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-7	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-7	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-12	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-12	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-12	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-12	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-12	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-12	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-12	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.15.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.15.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.15.0-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.15.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.5.2-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	12.3.0	Nvidia 드라이버와 Cuda 도구키트

구성 요소	버전	설명
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.285-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.285-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.285-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	442-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	442-amzn-0	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
trino-client	442-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.1-amzn-1	Spark 명령줄 클라이언트.
spark-history-server	3.5.1-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.1-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.1-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	24.06.1-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.16.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-10	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
tez-on-worker	0.10.2-amzn-10	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
zeppelin-server	0.11.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.1-amzn-1	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.1-amzn-1	ZooKeeper 명령줄 클라이언트.

7.3.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-7.3.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.

분류	설명	재구성 작업
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services SecondaryNamenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatchAgent service.

7.3.0 변경 로그

7.3.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2024-10-17	문서 게시	Amazon EMR 7.3.0 릴리스 정보 최초 게시됨
2024-10-16	배포 완료	모든 지원되는 리전에 Amazon EMR 7.3.0 전체 배포됨
2024-10-14	초기 릴리스	최초 상용 리전에 Amazon EMR 7.3.0 최초 배포됨

Amazon EMR 릴리스 7.2.0

7.2.0 지원 수명 주기

다음 표에서는 Amazon EMR 릴리스 7.2.0에 대해 지원되는 수명 주기 날짜를 설명합니다.

지원 단계	날짜	
초기 릴리스 날짜	2024년 7월 25일	
표준 지원 종료 시점	2026년 7월 24일	
지원 종료	2026년 7월 25일	
수명 종료	2027년 7월 25일	

권장 소프트웨어 및 구성에 따라 EMR 클러스터에 영향을 미치지 않는 중요한 상위의 일반적인 취약성 및 노출(CVE) 목록은 7.2.0 코어 엔진에 대해 알려진 일반적인 취약성 및 노출을 참조하세요. 7.2.0 릴리스에서 수정된 CVE 목록은 7.2.0 수정된 일반적인 취약성 및 노출을 참조하세요.

7.2.0 릴리스 구성 요소

Amazon EMR 7.2.0에는 다음 애플리케이션이 포함되어 있습니다.

코어 엔진:

AmazonCloudWatchAgent, Delta. Flink, Hadoop, HBase, HCatalog, Hive, Hudi, Livy, Iceberg, JupyterEnterpriseGateway, Phoenix, Spark, Tez, Trino, Zookeeper

추가 항목:

Hue, JupyterHub, MXNet, Oozie, Pig, Presto, Sqoop, TensorFlow, Zeppelin

포함된 패키지에 대한 자세한 내용은 Amazon EMR 7.x 릴리스의 애플리케이션 버전을 참조하세요. 코어 엔진 및 추가 항목에 대한 자세한 내용은 Amazon EMR 표준 지원을 참조하세요.

7.2.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS Java용 SDK	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.285	0.285	0.284	0.283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

7.2.0 일반적인 취약성 및 노출

이 페이지에는 Amazon EMR 7.2.0과 관련된 일반적인 취약성 및 노출(CVE)이 모두 나와 있습니다.

7.2.0 코어 엔진에 대한 알려진 일반적인 취약성 및 노출

다음 표에는 Amazon EMR 7.2.0의 권장 구성에서 실행되는 EMR 클러스터에 영향을 주지 않는 모든 CVE가 나와 있습니다. Amazon EMR은 업스트림 오픈 소스의 수정 사항 가용성에 종속되어 있으며, Amazon EMR이 수정 사항을 확인한 후 90일 이내에 코어 엔진 구성 요소의 일부로 안정적인 최신 버전을 제공합니다.

이 표는 2024년 9월 16일에 업데이트되었습니다.

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-40152	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40152
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-36478
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2015-5237	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-5237
CVE-2017-15095	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-17485

CVE ID	심각도	CVE 세부 정보 URL
CVE-2017-7525	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-7525
CVE-2018-11307	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-7489

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-10172	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-10172
CVE-2019-10202	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-10202
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17531

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-20330	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-14062	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-35490	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2020-8840	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2021-33813	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-33813

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-25194	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-34610	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34610

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-3635	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2012-0881	높음	https://nvd.nist.gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	높음	https://nvd.nist.gov/vuln/detail/CVE-2013-4002
CVE-2015-1832	심각	https://nvd.nist.gov/vuln/detail/CVE-2015-1832
CVE-2015-5237	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-5237
CVE-2018-11307	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14719

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-14720	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19360
CVE-2018-19361	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19362
CVE-2018-7489	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16335

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-16942	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2019-17267	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17531
CVE-2019-20330	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11113

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-11619	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2020-13936	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13936
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2020-35490	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2020-36186	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-8840	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2021-31684	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-31684
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-36364	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-36364
CVE-2022-37865	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-37865
CVE-2022-37866	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-37866
CVE-2022-39135	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-39135
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2022-46337	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2022-46751	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42889	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-42889
CVE-2022-46751	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-46751
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-20883	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-20883
CVE-2023-25194	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34455

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2023-5072	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-5072
CVE-2018-8768	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-8768
CVE-2021-32797	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-32797
CVE-2021-32798	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-32798
CVE-2022-24758	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-24758
CVE-2023-24580	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-24580
CVE-2023-36053	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-36053
CVE-2023-41164	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-41164
CVE-2023-43665	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43665

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-46695	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-46695
CVE-2024-24680	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-24680
CVE-2016-5018	심각	https://nvd.nist.gov/vuln/detail/CVE-2016-5018
CVE-2016-6796	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-6796
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2012-0881	높음	https://nvd.nist.gov/vuln/detail/CVE-2012-0881
CVE-2013-4002	높음	https://nvd.nist.gov/vuln/detail/CVE-2013-4002
CVE-2019-17566	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-17566
CVE-2019-17566	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-17566
CVE-2020-11988	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11988

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-33813	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-33813
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-34169	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-34169
CVE-2022-40146	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40146
CVE-2022-41704	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-41704
CVE-2022-42890	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42890
CVE-2022-44729	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-44729
CVE-2022-44729	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-44729
CVE-2022-44729	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-44729
CVE-2022-46337	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2023-4759	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-4759
CVE-2015-2156	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-2156
CVE-2017-18640	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-18640

CVE ID	심각도	CVE 세부 정보 URL
CVE-2017-18640	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-18640
CVE-2019-16869	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-16869
CVE-2019-20444	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20445
CVE-2021-37136	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37137
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-36478	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-36478
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2015-2080	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-2080
CVE-2015-4852	심각	https://nvd.nist.gov/vuln/detail/CVE-2015-4852
CVE-2015-6420	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-6420
CVE-2015-7501	심각	https://nvd.nist.gov/vuln/detail/CVE-2015-7501
CVE-2017-1000034	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-1000034
CVE-2017-12612	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-12612

CVE ID	심각도	CVE 세부 정보 URL
CVE-2017-15288	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-15288
CVE-2017-5637	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-5637
CVE-2017-7656	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-7656
CVE-2017-7657	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-7657
CVE-2017-7658	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-7658
CVE-2017-7687	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-7687
CVE-2017-9735	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-9735
CVE-2017-9790	높음	https://nvd.nist.gov/vuln/detail/CVE-2017-9790
CVE-2018-11793	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-11793
CVE-2018-11804	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-11804
CVE-2018-17190	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-17190
CVE-2018-8012	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-8012
CVE-2019-0204	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0204

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-10099	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-10099
CVE-2020-17533	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-17533
CVE-2020-9480	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9480
CVE-2021-28165	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-28165
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35517
CVE-2021-35517	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36090
CVE-2021-36090	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36090
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-37866	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-37866
CVE-2022-46751	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-46751
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2014-0114	높음	https://nvd.nist.gov/vuln/detail/CVE-2014-0114
CVE-2016-1000338	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-1000338

CVE ID	심각도	CVE 세부 정보 URL
CVE-2016-1000340	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-1000340
CVE-2016-1000342	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-1000342
CVE-2016-1000343	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-1000343
CVE-2016-1000344	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-1000344
CVE-2016-1000352	높음	https://nvd.nist.gov/vuln/detail/CVE-2016-1000352
CVE-2017-1000487	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-1000487
CVE-2017-12629	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-12629
CVE-2018-10936	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-10936
CVE-2018-3258	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-3258
CVE-2018-3827	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-3827
CVE-2018-3831	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-3831
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2019-10086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-10086

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-10086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-10086
CVE-2019-7611	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-7611
CVE-2020-13692	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13692
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2020-28491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-28491
CVE-2020-7692	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-7692
CVE-2021-22573	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-22573
CVE-2021-22573	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-22573
CVE-2021-25646	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-25646
CVE-2021-26291	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-26291
CVE-2021-26291	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-26291
CVE-2021-26919	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-26919
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516
CVE-2021-35517	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36090
CVE-2022-1471	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-1471
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-25857	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25857
CVE-2022-31197	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-31197
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-33980	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-33980
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-4244	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-4244
CVE-2022-45688	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45688
CVE-2023-1370	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1370
CVE-2023-1428	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1428
CVE-2023-25194	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-25194
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-31418	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-31418
CVE-2023-32731	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-32731

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-38647
CVE-2023-39410	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-39410
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642
CVE-2023-44981	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-44981
CVE-2023-46120	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-46120
CVE-2023-5072	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-5072
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-1597

CVE ID	심각도	CVE 세부 정보 URL
CVE-2024-1597	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-1597
CVE-2021-35958	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-35958
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2021-34538	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-34538
CVE-2022-45688	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45688
CVE-2022-46337	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-46337
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-3635	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2023-5072	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-5072
CVE-2024-21634	높음	https://nvd.nist.gov/vuln/detail/CVE-2024-21634
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-35517	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35517
CVE-2021-36090	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-36090
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2014-0114	높음	https://nvd.nist.gov/vuln/detail/CVE-2014-0114
CVE-2015-2156	높음	https://nvd.nist.gov/vuln/detail/CVE-2015-2156
CVE-2018-10054	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-10054
CVE-2019-0205	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-0205
CVE-2019-10086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-10086
CVE-2019-10086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-10086

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-16869	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-16869
CVE-2019-17195	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17195
CVE-2019-20444	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20444
CVE-2019-20445	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20445
CVE-2020-13949	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-13949
CVE-2020-9492	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-9492
CVE-2021-37136	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37136
CVE-2021-37137	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37137
CVE-2021-37404	심각	https://nvd.nist.gov/vuln/detail/CVE-2021-37404
CVE-2022-25168	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-25168
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-26612	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-26612

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-3171	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3171
CVE-2022-3509	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3509
CVE-2022-3510	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-3510
CVE-2022-40152	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40152
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-34453	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34453
CVE-2023-34454	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34454
CVE-2023-34455	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-34455
CVE-2023-3635	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2023-3635	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-3635
CVE-2023-38647	심각	https://nvd.nist.gov/vuln/detail/CVE-2023-38647
CVE-2023-43642	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-43642

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481
CVE-2024-1597	심각	https://nvd.nist.gov/vuln/detail/CVE-2024-1597
CVE-2017-15095	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-15095
CVE-2017-17485	심각	https://nvd.nist.gov/vuln/detail/CVE-2017-17485
CVE-2018-11307	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-11307
CVE-2018-12022	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12022
CVE-2018-12023	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-12023
CVE-2018-14718	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14718
CVE-2018-14719	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14719
CVE-2018-14720	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14720
CVE-2018-14721	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-14721
CVE-2018-19360	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19360

CVE ID	심각도	CVE 세부 정보 URL
CVE-2018-19361	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19361
CVE-2018-19362	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-19362
CVE-2018-5968	높음	https://nvd.nist.gov/vuln/detail/CVE-2018-5968
CVE-2018-7489	심각	https://nvd.nist.gov/vuln/detail/CVE-2018-7489
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086
CVE-2019-12086	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12086
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2019-14379	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14379
CVE-2019-14439	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2019-14439	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-14439
CVE-2019-14540	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2019-14540	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14540
CVE-2019-14892	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14892

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-14892	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14892
CVE-2019-14893	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-14893
CVE-2019-16335	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16335
CVE-2019-16335	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16335
CVE-2019-16942	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2019-16942	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16942
CVE-2019-16943	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2019-16943	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-16943
CVE-2019-17195	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17195
CVE-2019-17267	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2019-17267	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17267
CVE-2019-17531	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17531
CVE-2019-17531	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-17531

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-20330	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2019-20330	심각	https://nvd.nist.gov/vuln/detail/CVE-2019-20330
CVE-2020-10650	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-10650	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10650
CVE-2020-10672	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-10672	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10672
CVE-2020-10673	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10673
CVE-2020-10673	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10673
CVE-2020-10968	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10968
CVE-2020-10968	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10968
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2020-10969	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-10969
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-11111	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11111
CVE-2020-11112	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11112
CVE-2020-11113	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11113
CVE-2020-11619	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11619
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2020-11620	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-11620
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2020-14060	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14060
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061
CVE-2020-14061	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14061
CVE-2020-14062	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-14062	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14062
CVE-2020-14195	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14195

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-14195	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-14195
CVE-2020-24616	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-24616	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24616
CVE-2020-24750	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-24750	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-24750
CVE-2020-25649	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2020-25649	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2020-25649	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-25649
CVE-2020-35211	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35211
CVE-2020-35490	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2020-35490	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35490
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491
CVE-2020-35491	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35491
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2020-35728	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-35728
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2020-36179	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36179
CVE-2020-36180	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2020-36180	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2020-36180	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36180
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36181	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36181
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2020-36182	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36182
CVE-2020-36183	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2020-36183	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2020-36183	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36183
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2020-36184	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36184
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185
CVE-2020-36185	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36185

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36186	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2020-36186	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36186
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-36187	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36187
CVE-2020-36188	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-36188	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-36188	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36188
CVE-2020-36189	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2020-36189	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36189
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518

CVE ID	심각도	CVE 세부 정보 URL
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2020-36518	높음	https://nvd.nist.gov/vuln/detail/CVE-2020-36518
CVE-2020-7692	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-7692
CVE-2020-8840	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-8840	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-8840
CVE-2020-9546	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9546
CVE-2020-9546	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9546
CVE-2020-9547	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-9547	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9547
CVE-2020-9548	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9548
CVE-2020-9548	심각	https://nvd.nist.gov/vuln/detail/CVE-2020-9548
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190

CVE ID	심각도	CVE 세부 정보 URL
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2021-20190	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-20190
CVE-2021-22573	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-22573
CVE-2021-33036	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-33036
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515
CVE-2021-35515	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35515
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516
CVE-2021-35516	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-35516
CVE-2021-37714	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37714
CVE-2021-37714	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-37714
CVE-2021-46877	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-46877

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-25647	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-25647
CVE-2022-32532	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-32532
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-40664	심각	https://nvd.nist.gov/vuln/detail/CVE-2022-40664
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42003	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42003
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004
CVE-2022-42004	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-42004

CVE ID	심각도	CVE 세부 정보 URL
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45688	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45688
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2023-2976	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-2976
CVE-2022-40149	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40149
CVE-2022-40150	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-40150
CVE-2022-45685	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45685
CVE-2022-45693	높음	https://nvd.nist.gov/vuln/detail/CVE-2022-45693
CVE-2023-1436	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-1436
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378
CVE-2023-6378	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6378

CVE ID	심각도	CVE 세부 정보 URL
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481
CVE-2023-6481	높음	https://nvd.nist.gov/vuln/detail/CVE-2023-6481

7.2.0 코어 엔진에 대한 일반적인 취약성 및 노출 수정됨

다음 표에는 Amazon EMR 7.2.0에서 수정된 모든 CVE가 나와 있습니다.

CVE ID	심각도	CVE 세부 정보 URL
CVE-2019-12402	높음	https://nvd.nist.gov/vuln/detail/CVE-2019-12402
CVE-2021-41561	높음	https://nvd.nist.gov/vuln/detail/CVE-2021-41561

7.2.0 릴리스 구성 요소 세부 정보

Amazon EMR 7.2.0의 추가 항목 패키지에 대한 자세한 내용은 다음 표를 참조하세요.

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.4-1.amzn2023
AmazonCloudWatchAgent	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
AmazonCloudWatchAgent	Core	s3-dist-cp	추가 항목	2.32.0-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
AmazonCloudWatchAgent	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
AmazonCloudWatchAgent	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Flink	Core	hadoop-client	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-kms	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn	추가 항목	3.3.6.amzn.4-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Flink	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Flink	Core	hudi	추가 항목	0.14.1.amzn.1-1.amzn2023
Flink	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Flink	Core	hadoop	추가 항목	3.3.6.amzn.4-1.amzn2023
HBase	Core	s3-dist-cp	추가 항목	2.32.0-1.amzn2023
HCatalog	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
HCatalog	Core	mariadb-connector-java	추가 항목	2.7.2-1
Hive	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.4-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Hive	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Hive	Core	hive-hbase	추가 항목	3.1.3.amzn.11-1.amzn2023
Hive	Core	mariadb-connector-java	추가 항목	2.7.2-1
Hive	Core	s3-dist-cp	추가 항목	2.32.0-1.amzn2023
피닉스	Core	s3-dist-cp	추가 항목	2.32.0-1.amzn2023
Spark	Core	spark-datanucleus	추가 항목	3.5.1.amzn.0-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.4-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Tez의 Hive 실행 시간 비교	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	bigtop-utils	추가 항목	1.2.0-1.amzn2023
Trino	Core	hadoop-client	추가 항목	3.3.6.amzn.4-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-kms	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.4-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hive	추가 항목	3.1.3.amzn.11-1.amzn2023
Trino	Core	hive-hcatalog-server	추가 항목	3.1.3.amzn.11-1.amzn2023
Trino	Core	mariadb-connector-java	추가 항목	2.7.2-1
Trino	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023
Trino	Core	zookeeper	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	zookeeper-server	추가 항목	3.9.1.amzn.1-1.amzn2023
Trino	Core	bigtop-groovy	추가 항목	2.5.4-1.amzn2023
Trino	Core	bigtop-jsvc	추가 항목	1.2.4-1.amzn2023
Trino	Core	hadoop	추가 항목	3.3.6.amzn.4-1.amzn2023
Trino	Core	hadoop-lzo	추가 항목	0.4.19-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Trino	Core	hive-hcatalog	추가 항목	3.1.3.amzn.11-1.amzn2023
Trino	Core	hive-jdbc	추가 항목	3.1.3.amzn.11-1.amzn2023
Zookeeper	Core	emrfs	추가 항목	2.63.0-1.amzn2023
Zookeeper	Core	hadoop-client	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-hdfs	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-hdfs-datanode	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-hdfs-journalnode	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-hdfs-namenode	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-httpfs	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-kms	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-mapreduce	추가 항목	3.3.6.amzn.4-1.amzn2023

Application	애플리케이션 유형	Package	패키지 유형	패키지 버전
Zookeeper	Core	hadoop-mapreduce-historyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-yarn	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-yarn-nodemanager	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-yarn-proxyserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-yarn-resourcemanager	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	hadoop-yarn-timelineserver	추가 항목	3.3.6.amzn.4-1.amzn2023
Zookeeper	Core	ranger-kms	추가 항목	2.0.0-1.amzn2023

7.2.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.2.0에 대한 정보가 포함됩니다. 변경 사항은 7.2.0에 관련됩니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.2.0 application upgrades include Iceberg 1.5.0-amzn-0 and Delta 3.1.0.
- Amazon EMR은 Amazon S3 Express One Zone 스토리지 클래스에서 HBase, Flink 및 Hive와 같은 다른 애플리케이션을 사용할 수 있도록 지원을 추가합니다.

- 이 릴리스에서는 복원된 객체를 읽을 수 있는 기능을 추가하므로, S3A 프로토콜로 S3 위치에서 Glacier 객체를 읽을 수 있습니다. 이 기능은 Spark, Flink 및 Hive에서 작동합니다.
- 관리형 조정을 사용하는 노드 레이블 - 시장 유형 또는 노드 유형에 따라 인스턴스에 노드 레이블을 사용하여 Amazon EMR에서 자동 조정을 개선합니다. 자세한 내용은 [Amazon EMR에서 관리형 조정 사용](#)을 참조하세요.

알려진 문제

- Python 3.11은 EMR Studio에서 지원되지 않습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 이 릴리스는 내부 단계 정리 작업 중에 발생할 수 있는 교착 문제를 수정합니다. 이 작업은 EMR 클러스터에서 완료할 때 단계의 수명 주기를 관리합니다. 이 문제는 단계 작업 및 조정과 같은 중요한 Amazon EMR 작업에 영향을 미칩니다.
- 이 릴리스는 특정 기존 로그 파일이 있는 사용자 지정 AMI를 사용하는 사용자 지정 클러스터로 인해 Amazon EMR 로그 관리 대몬이 실패할 수 있는 문제를 해결합니다.
- Amazon EMR 7.2.0은 클러스터 관리 및 모니터링 활동을 담당하는 Amazon EMR 데몬을 AWS SDK v1에서 v2로 업그레이드합니다.

- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용](#)을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2일	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250124	6.1.124-134.200.amzn2023	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250110	6.1.119-129.201.amzn2023	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 11월 21일	6.1.115-126.197.amzn2023	2024년 12월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 10월 31.0일	6.1.112-124.190.amzn2023	2024년 11월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 1001.1일	6.1.109-118.189.amzn2023	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 8월 19일	6.1.102-111.182.amzn2023	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 30일	6.1.97-104.177.amzn2023	2024년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 22일	6.1.97-104.177.amzn2023	2024년 7월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 8일	6.1.96-102.177.amzn2023	2024년 7월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

7.2.0 기본 Java 버전

Amazon EMR 릴리스 7.2 이상에서는 Apache Livy를 제외한 Corretto 17(JDK 17)을 지원하는 애플리케이션에 대해 기본적으로 Amazon Corretto 17(OpenJDK 기반)이 함께 제공됩니다.

다음 표에서는 Amazon EMR 7.2.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	21, 17
Zeppelin	8
Pig	8
Zookeeper	8

7.2.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.1.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.4.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.11.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.15.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.16.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.32.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.11.0	EMR S3Select 커넥터
emr-wal-cli	1.2.2	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.63.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.18.1-amzn-1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.18.1-amzn-1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.3.6-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.6-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.6-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.6-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.6-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.6-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.6-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.6-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.6-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.6-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.6-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.4.17-amzn-6	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-6	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-6	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-6	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-6	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-6	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-11	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-11	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-11	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-11	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-11	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-11	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-11	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.14.1-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.14.1-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.14.1-amzn-1	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.14.1-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.5.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.

구성 요소	버전	설명
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.3	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.285-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.285-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.285-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	436-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
trino-worker	436-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	436-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.1-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.5.1-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.1-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.1-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	24.02.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-9	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
tez-on-worker	0.10.2-amzn-9	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.1-amzn-1	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.1-amzn-1	ZooKeeper 명령줄 클라이언트.

7.2.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-7.2.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.

분류	설명	재구성 작업
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logba ck.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.	Restarts Zookeeper server.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.	Restarts the CloudWatchAgent service.

7.2.0 변경 로그

7.2.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2024-07-25	문서 게시	Amazon EMR 7.2.0 릴리스 정보 최초 게시됨
2024-07-16	초기 릴리스	최초 상용 리전에 Amazon EMR 7.2.0 최초 배포됨

Amazon EMR 릴리스 7.1.0

7.1.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전

- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS Java용 SDK	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.285	0.285	0.284	0.283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

7.1.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.1.0에 대한 정보가 포함됩니다. 변경 사항은 7.1.0에 관련됩니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.1.0 application upgrades include Livy 0.8.0, Trino 435, and ZooKeeper 3.9.1.

- 비정상 노드 교체 - Amazon EMR 7.1.0 이상에서는 비정상 노드 교체가 기본적으로 활성화되므로 Amazon EMR이 비정상 노드를 정상적으로 교체합니다. Amazon EMR 릴리스 7.0.0 이하의 기존 워크플로에 영향을 주지 않도록 클러스터에서 종료 방지를 활성화한 경우 비정상 노드 교체는 비활성화됩니다.
- CloudWatch 에이전트 - Amazon EMR 구성 API를 사용하여 추가 시스템 지표를 사용하고, 애플리케이션 지표를 추가하며, 지표 대상을 변경하도록 CloudWatch 에이전트를 구성합니다.

알려진 문제

- JupyterEnterpriseGateway(JEG) 애플리케이션은이 릴리스에서 Python 3.11을 지원하지 않습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Amazon EMR 7.1.0은 기본적으로 Python 3.9를 지원하지만 Amazon EMR 7.1.0의 Livy 0.8.0 및 Spark는 Python 3.11을 지원합니다.
- 이 릴리스에서는 PySpark를 Python 버전 3.11과 함께 사용할 때 한 번에 하나씩 각 줄을 실행해야 하던 문제를 수정합니다.
- Zeppelin 업그레이드 - Amazon EMR 7.1.0에는 Zeppelin을 AWS SDK for Java v2로 업그레이드하는 작업이 포함되어 있습니다. 이 업그레이드를 통해 Zeppelin S3 노트북은 사용자 지정 암호화 구성 요소 제공업체를 수락할 수 있습니다. AWS SDK for Java v2는 EncryptionMaterialsProvider 인터페이스를 제거합니다. Amazon EMR 7.1.0으로 업그레이드하는 경우 사용자 지정 암호화를 사용하려면 Keyring 인터페이스를 구현해야 합니다. 키링 인터페이스를 구현하는 방법의 예제는 KmsKeyring.java를 참조하세요.
- Amazon EMR 릴리스 7.1.0으로 업그레이드하는 경우 AES/GCM/NoPadding 의 AES 알고리즘을 사용하여 키를 생성하도록 로컬 디스크 암호화에 대한 사용자 지정 키 제공업체를 변경합니다. 알

고리즘을 업데이트하지 않으면 `Local disk encryption failed on master instance (i-123456789) due to internal error.` 오류와 함께 클러스터 생성이 실패할 수 있습니다. 사용자 지정 키 제공업체 생성에 대한 자세한 내용은 사용자 지정 키 제공업체 생성을 참조하세요.

- Amazon EMR 7.1.0은 파일 핸들이 열려 있는 파일에 대한 로그 잘림 로직을 개선하여 디스크 공간 부족 조건에서 노드의 복원력을 개선합니다.
- 이 릴리스는 인코딩 및 디코딩 로직을 개선하여 노드를 재시작할 때 Amazon EMR 대몬의 읽기 및 쓰기 파일을 사용하여 데이터 손상 및 노드 실패 위험을 최소화합니다.

- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2일	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250124	6.1.124-134.200.amzn2023	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250110	6.1.119-129.201.amzn2023	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 11월 21일	6.1.115-126.197.amzn2023	2024년 12월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 10월 31.0일	6.1.112-124.190.amzn2023	2024년 11월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 1001.1일	6.1.109-118.189.amzn2023	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 8월 19일	6.1.102-111.182.amzn2023	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 30일	6.1.97-104.177.amzn2023	2024년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 22일	6.1.97-104.177.amzn2023	2024년 7월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 8일	6.1.96-102.177.amzn2023	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 3월 2024년 219.0	6.1.77-99.164.amzn2023	2024년 5월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

7.1.0 기본 Java 버전

Amazon EMR 릴리스 7.1 이상에서는 Apache Livy를 제외한 Corretto 17(JDK 17)을 지원하는 애플리케이션에 대해 기본적으로 Amazon Corretto 17(OpenJDK 기반)이 함께 제공됩니다.

다음 표에서는 Amazon EMR 7.1.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	17, 11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

7.1.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.0.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
emr-amazon-cloudwatch-agent	1.300032.2-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.10.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.14.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.31.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.10.0	EMR S3Select 커넥터
emr-wal-cli	1.2.2	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.62.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.18.1-amzn-0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.18.1-amzn-0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
hadoop-client	3.3.6-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.6-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.6-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.6-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.6-amzn-3	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.6-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.6-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.6-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.6-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.6-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.6-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	2.4.17-amzn-5	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-5	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-5	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-5	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-5	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-5	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-10	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-10	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-10	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-10	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-10	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-10	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.3-amzn-10	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.14.1-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.14.1-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.14.1-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.14.1-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.4.3-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.8.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.

구성 요소	버전	설명
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.3	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.284-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.284-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.284-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	435-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
trino-worker	435-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	435-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.0-amzn-1	Spark 명령줄 클라이언트.
spark-history-server	3.5.0-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.0-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.0-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.10.0-amzn-1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-8	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
tez-on-worker	0.10.2-amzn-8	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.9.1-amzn-0	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.9.1-amzn-0	ZooKeeper 명령줄 클라이언트.

7.1.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-7.1.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
docker-conf	Docker 관련 설정을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-env	HDFS 환경의 값을 변경합니다.

분류	설명
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.

분류	설명
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.

분류	설명
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.

분류	설명
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.

분류	설명
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.

분류	설명
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-logback	ZooKeeper의 logback.xml 파일에서 값을 변경합니다.
emr-metrics	이 노드에 대한 emr 지표 설정을 변경합니다.

7.1.0 변경 로그

7.1.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2024-05-10	문서 게시	Amazon EMR 7.1.0 릴리스 정보 최초 게시됨

날짜	이벤트	설명
2024-05-08	배포 완료	모든 지원되는 리전에 Amazon EMR 7.1.0 전체 배포됨
2024-04-23	초기 릴리스	최초 상용 리전에 Amazon EMR 7.1.0 최초 배포됨

Amazon EMR 릴리스 7.0.0

7.0.0 애플리케이션 버전

이 릴리스에는 AmazonCloudWatchAgent, , Delta, Flink, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Presto, PigSparkSqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AWS Java용 SDK	2.25.70, 1.12.747	2.23.18, 1.12.705	2.23.18, 1.12.656	2.20.160-amzn-0, 1.12.569
Python	3.9, 3.11	3.9, 3.11	3.9, 3.11	3.9
Scala	2.12.18	2.12.17	2.12.17	2.12.17

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
AmazonCloudWatchAgent	1.300032.2	1.300032.2	1.300032.2	1.300031.1
Delta	3.2.0-amzn-0	3.1.0	3.0.0	3.0.0
Flink	1.18.1	1.18.1	1.18.1	1.18.0
Ganglia	-	-	-	-
HBase	2.4.17	2.4.17	2.4.17	2.4.17
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.6	3.3.6	3.3.6
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.15.0-amzn-0	0.14.1-amzn-1	0.14.1-amzn-0	0.14.0-amzn-1
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.5.2-amzn-0	1.5.0-amzn-0	1.4.3-amzn-0	1.4.2-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.5.0
Livy	0.8.0	0.8.0	0.8.0	0.7.1
MXNet	-	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0

	emr-7.3.0	emr-7.2.0	emr-7.1.0	emr-7.0.0
Presto	0.285	0.285	0.284	0.283
Spark	3.5.1	3.5.1	3.5.0	3.5.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.16.1	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	442	436	435	426
Zeppelin	0.11.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.9.1	3.9.1	3.9.1	3.5.10

7.0.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 7.0.0에 대한 정보가 포함됩니다. 변경 사항은 6.15.0에 관련됩니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.0.0 application upgrades include Python 3.9, Spark 3.5, Flink 1.18, and Delta 3.0. 또한 이번 릴리스에서는 Amazon CloudWatch 에이전트 애플리케이션에 대한 지원이 추가되고 Ganglia에 대한 지원이 제거됩니다.
- Amazon Corretto 17 - Amazon EMR 릴리스 7.0 이상에서는 Apache Livy를 제외한 Corretto 17(JDK 17)을 지원하는 애플리케이션에 대해 기본적으로 Amazon Corretto 17(OpenJDK 기반)이 함께 제공됩니다. 이번 릴리스에서 지원되는 애플리케이션의 JDK 버전에 대한 자세한 내용을 확인하려면 7.0.0 기본 Java 버전 섹션을 참조하세요.
- Amazon Linux 2023 – 7.0 릴리스를 사용하면 이제 Amazon EMR 클러스터가 기본적으로 AL2023에서 실행됩니다. 이러한 변경이 기본 AMI 버전에 어떤 영향을 미치는지에 대한 자세한 내용을 확인하려면 Amazon EMR 관리 안내서의 소프트웨어 업데이트 고려 사항을 참조하세요. 또한 AL2023에서 Python 2.7이 제거되었으므로 이제 Python이 필요한 모든 구성 요소는 Python 3을 이용하여 작성해야 합니다.

- **s3a**가 포함된 S3 on Outpost – 이제 Amazon EMR에서는 `s3a` 파일 시스템이 포함된 Amazon S3 on Outposts 버킷을 지원합니다. Outposts의 Amazon S3에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 [S3 on Outposts이란](https://docs.aws.amazon.com/AmazonS3/latest/userguide/S3onOutposts.html)을 참조하세요.

알려진 문제

- PySpark를 Python 버전 3.10 이상과 함께 사용하는 경우 한 번에 두 개 이상의 줄을 실행할 수 없습니다. 한 번에 한 줄씩 실행해야 합니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정합니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 클러스터에 있는 모든 인스턴스의 상태를 유지하기 위해 Amazon EMR은 Apache YARN, Apache HDFS 및 Kerberos와 통합됩니다. 7.0에서는 이러한 통합 기능을 개선하여 확장 또는 기타 작업으로 인해 종료된 인스턴스의 상태를 안정적으로 제거할 수 있도록 했습니다. 이러한 노력은 관리형 확장이 활성화된 오래된 클러스터의 경우, 자주 규모를 조정하고 수명 기간 동안 종료된 인스턴스가 수천 개 누적되기 때문에 특히 중요합니다.
- 이번 릴리스는 Kerberos 구성을 개선하여 AES 기반 암호만 지원하도록 했습니다. AES 기반이 아닌 암호를 사용하는 Kerberos KDC는 Amazon EMR 릴리스 7.0.0 이상에서 실행되는 EMR 클러스터에서 더 이상 지원되지 않습니다. AES 기반 암호는 클러스터에 가장 강력한 보안을 제공합니다.
- AWS SDK 2.x 마이그레이션의 일환으로 Amazon EMR 7.0에는 호환성을 위해 Spark Kinesis 커넥터에 대한 업데이트가 포함되어 있습니다. 이 업데이트는 커뮤니티 버전의 Apache Spark에서는 이용할 수 없습니다. 7.0 미만의 Amazon EMR 릴리스에서 Spark Kinesis 커넥터를 사용하는 경우, 워크로드를 Amazon EMR 7.0으로 마이그레이션하기 전에 애플리케이션 코드를 마이그레이션하여 SDK 2.x에서 실행되도록 해야 합니다. 자세한 내용은 [Spark Kinesis 커넥터를 Amazon EMR 7.0용 SDK 2.x로 마이그레이션](#) 단원을 참조하십시오.

- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 5월 12일	6.1.134-152.225.amzn2023	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 428.1일	6.1.134-150.224.amzn2023	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 4월 14일	6.1.132-147.221.amzn2023	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250124	6.1.124-134.200.amzn2023	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 20250110	6.1.119-129.201.amzn2023	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 11월 21일	6.1.115-126.197.amzn2023	2024년 12월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2024년 1031.0	6.1.112-124.190.amzn2023	2024년 11월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 1001.1일	6.1.109-118.189.amzn2023	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 8월 19일	6.1.102-111.182.amzn2023	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 30일	6.1.97-104.177.amzn2023	2024년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 22일	6.1.97-104.177.amzn2023	2024년 7월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 5월 2024년 7월 8일	6.1.96-102.177.amzn2023	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 3월 2024년 304.0	6.1.79-99.164.amzn2023	2024년 3월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 3월 2024년 219.0	6.1.77-99.164.amzn2023	2024년 3월 1일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 3월 2024년 205.0	6.1.75-99.163.amzn2023	2024년 2월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 3월 2024년 122.0	6.1.72-96.166.amzn2023	2024년 2월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 3월 240108.(	6.1.72-96.166.amzn2023	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), 캐나다 서부(캘거리), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 3월 2023년 1211.4	6.1.66-91.160.amzn2023	2023년 12월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

7.0.0 기본 Java 버전

Amazon EMR 릴리스 7.0 이상에서는 Apache Livy를 제외한 Corretto 17(JDK 17)을 지원하는 애플리케이션에 대해 기본적으로 Amazon Corretto 17(OpenJDK 기반)이 함께 제공됩니다.

다음 표에서는 Amazon EMR 7.0.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	11, 8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

7.0.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
adot-java-agent	1.31.0	애플리케이션 대몬(daemon)에서 지표를 수집하는 Java 에이전트입니다.
delta	3.0.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
delta-standalone-connectors	0.6.0	Delta 커넥터는 Delta Lake를 Flink, Hive, Presto와 같은 엔진과 통합하기 위한 다양한 런타임을 제공합니다.
emr-amazon-cloudwatch-agent	1.300031.1-amzn-0	Amazon EC2 인스턴스에서 내부 시스템 수준 지표와 사용자 지정 애플리케이션 지표를 수집하는 애플리케이션입니다.
emr-ddb	5.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.9.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.13.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.30.0	Amazon S3용으로 최적화된 분산 복사 애플리케이션.
emr-s3-select	2.9.0	EMR S3Select 커넥터
emr-wal-cli	1.2.1	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.61.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.18.0-amzn-0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
flink-jobmanager-config	1.18.0-amzn-0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
hadoop-client	3.3.6-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.6-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.6-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.6-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.6-amzn-2	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.6-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.6-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.6-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.6-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.6-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	3.3.6-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.17-amzn-4	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-4	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-4	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-4	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-4	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-4	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-9	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-9	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-9	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-9	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-9	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	3.1.3-amzn-9	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-9	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.14.0-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.14.0-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.14.0-amzn-1	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.14.0-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.4.2-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.3	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.283-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.283-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.283-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.

구성 요소	버전	설명
trino-coordinator	426-amzn-1	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	426-amzn-1	여러 쿼리를 실행하는 서비스.
trino-client	426-amzn-1	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.3.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.5.0-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.5.0-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.5.0-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.5.0-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.10.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-7	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-7	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.58	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

7.0.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-7.0.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.

분류	설명
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
docker-conf	Docker 관련 설정을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-env	HDFS 환경의 값을 변경합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.

분류	설명
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.

분류	설명
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.

분류	설명
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.

분류	설명
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.

분류	설명
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

7.0.0 변경 로그

7.0.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-12-29	배포 완료	모든 지원되는 리전에 Amazon EMR 7.0.0 전체 배포됨
2023-12-28	문서 게시	Amazon EMR 7.0.0 릴리스 정보 최초 게시됨
2023-12-19	초기 릴리스	최초 상용 리전에 Amazon EMR 7.0.0 최초 배포됨

Amazon EMR 6.x 릴리스 버전

이 섹션에는 각 Amazon EMR 6.x 릴리스 버전에 제공되는 애플리케이션 버전, 릴리스 정보, 구성 요소 버전 및 구성 분류가 포함되어 있습니다.

클러스터를 시작할 때 Amazon EMR의 여러 릴리스 중에서 선택할 수 있습니다. 이를 통해 호환성 요구 사항에 맞는 애플리케이션 버전을 테스트하고 사용할 수 있습니다. 릴리스 레이블과 함께 릴리스 번호를 지정합니다. 릴리스 레이블은 `emr-x.x.x` 형식입니다. 예를 들어 `emr-7.9.0`입니다.

새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

모든 Amazon EMR 6.x 릴리스에서 애플리케이션 버전의 전체 테이블은 Amazon EMR 6.x 릴리스의 애플리케이션 버전 섹션을 참조하세요.

주제

Amazon EMR 6.x 릴리스의 애플리케이션 버전

각 Amazon EMR 6.x 릴리스에서 사용할 수 있는 애플리케이션 버전을 표시한 전체 테이블을 보려면 브라우저에서 Amazon EMR 6.x 릴리스의 애플리케이션 버전을 엽니다.

Amazon EMR 릴리스 6.15.0

6.15.0 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, HudiIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Presto, PigSparkSqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
AWS Java용 SDK	2.20.160-amzn-0, 1.12.569	1.12.543	1.12.513	1.12.490
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.17	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.4.0
Flink	1.17.1	1.17.1	1.17.0	1.17.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.17	2.4.17

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.6	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.14.0-amzn-0	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.4.0-amzn-0	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0
JupyterEn terpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.283	0.281	0.281	0.281
Spark	3.4.1	3.4.1	3.4.1	3.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2

	emr-6.15.0	emr-6.14.0	emr-6.13.0	emr-6.12.0
Trino (PrestoSQL)	426	422	414	414
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.15.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.15.0에 대한 정보가 포함됩니다. 변경 사항은 6.14.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.15.0 변경 로그 섹션을 참조하세요.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.
- EC2에서 실행되는 EMR 클러스터의 실행 속도 향상 – 이제 EC2 클러스터에서 Amazon EMR을 실행하는 속도가 최대 35% 빨라졌습니다. 이번 개선 덕분에 고객은 대부분 대략 5분 만에 클러스터를 시작할 수 있게 되었습니다.
- EMR Studio용 CodeWhisperer – 이제 JupyterLab에서 코드를 작성할 때 Amazon CodeWhisperer와 Amazon EMR Studio를 함께 사용하면 실시간으로 권장 사항을 받을 수 있습니다. CodeWhisperer에서는 설명을 작성하고, 코드 한 줄을 완성하고, 라인별 권장 사항을 만들고, 완전한 형식의 함수를 생성할 수 있습니다.
- Flink를 사용한 더욱 빠른 작업 재시작 시간 – Amazon EMR 6.15.0 이상을 사용하면 태스크 복구 또는 규모 조정 작업 중에 작업 재시작 시간을 단축하기 위해 Apache Flink에 대해 이제 몇 가지 새로운 메커니즘을 사용할 수 있습니다. 이는 실행 그래프의 복구 및 재시작 속도를 최적화하여 작업 안정성을 향상시킵니다.
- 개방형 테이블 형식에 대한 테이블 수준 및 세분화된 액세스 제어 - Amazon EMR 6.15.0 이상에서는 Glue 데이터 카탈로그의 데이터에 액세스하는 EC2 클러스터의 Amazon EMR에서 Spark 작업을 실행할 때 AWS AWS Lake Formation 를 사용하여 Hudi, Iceberg 또는 Delta Lake 기반 테이블에 테이블, 행, 열 및 셀 수준 권한을 적용할 수 있습니다.
- Hadoop 업그레이드 - Amazon EMR 6.15.0에는 Apache Hadoop의 버전 3.3.6으로의 업그레이드가 포함되어 있습니다. Hadoop 3.3.6은 Amazon EMR 6.15가 배포되었을 시점에 최신 버전이었으

며 Apache에서 2023년 6월에 릴리스했습니다. Amazon EMR의 이전 릴리스(6.9.0~6.14.x) 에서는 Hadoop 3.3.3이 사용되었습니다.

업그레이드에는 재구성 가능한 데이터노드 파라미터, 모든 라이브 데이터노드에서 대량 재구성 작업을 시작하는 `DFSAdmin` 옵션, 탐색이 많은 리더가 읽을 여러 범위를 지정할 수 있도록 하는 벡터 API 등 수백 가지 개선 사항과 수정 사항이 포함되어 있습니다. Hadoop 3.3.6에서는 WAL(미리 쓰기 로그)에 대한 HDFS API 및 의미 체계에 대한 지원도 추가되어 HBase가 다른 스토리지 시스템 구현에서 실행될 수 있습니다. 자세한 내용을 확인하려면Apache Hadoop 하둡 설명서에서 버전 3.3.4, 3.3.5 및 3.3.6의 변경 로그를 참조하세요.

- AWS SDK for Java 버전 2 지원 - 애플리케이션이 v2를 지원하는 경우 Amazon EMR 6.15.0 애플리케이션은 AWS SDK for Java 버전 1.12.569 또는 2.20.160을 사용할 수 있습니다. Java 2.x용 AWS SDK는 버전 1.x 코드 베이스의 주요 재작성입니다. Java 8+에 토대를 두고 있으며, 요청이 많았던 기능들을 몇 가지 추가했습니다. 여기에는 비차단 I/O에 대한 지원과 런타임에 다른 HTTP 구현을 연결하는 기능이 포함됩니다. Java용 SDK v1에서 v2로의 마이그레이션 가이드를 포함한 자세한 내용을 확인하려면 AWS Java용 SDK, 버전 2 가이드를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 고가용성 EMR 클러스터를 개선하기 위해 이번 릴리스에서는 IPv6 엔드포인트를 사용하는 로컬 호스트의 Amazon EMR 대몬(daemon)에 대한 연결을 지원합니다.
- 이번 릴리스는 고가용성 클러스터의 모든 프라이머리 노드에 프로비저닝된 ZooKeeper와의 통신을 위한 TLS 1.2가 지원합니다.
- 이번 릴리스에서는 로그 파일이 범위를 초과하여 클러스터 작업을 방해하는 시나리오를 최소화하기 위해 프라이머리 노드에서 유지되는 ZooKeeper 트랜잭션 로그 파일의 관리를 개선합니다.
- 이번 릴리스는 고가용성 EMR 클러스터에 대한 노드 내 통신의 복원력을 강화합니다. 이러한 개선 사항은 부트스트랩 작업 실패나 클러스터 시작 실패 가능성을 줄입니다.

- Amazon EMR 6.15.0의 Tez는 Tez 그룹화된 분할에서 입력 분할을 비동기적으로 열도록 지정할 수 있는 구성을 도입합니다. 그 결과 단일 Tez 그룹화된 분할에 많은 수의 입력 분할이 있을 경우 읽기 쿼리 성능이 더욱 빨라집니다. 자세한 내용을 확인하려면 Tez 비동기식 분할 열기를 참조하세요.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 13일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

6.15.0 기본 Java 버전

Amazon EMR 릴리스 6.12.0 이상에서는 기본적으로 Amazon Corretto 8을 사용하는 모든 애플리케이션을 지원합니다(단, Trino는 예외임). Trino와 관련하여 Amazon EMR은 Amazon EMR 릴리스 6.9.0부터 Amazon Corretto 17을 기본적으로 지원합니다. 또한 Amazon EMR은 Amazon Corretto 11 및 17을 사용하는 일부 애플리케이션도 지원합니다. 이 애플리케이션은 다음 테이블에 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Amazon EMR은 Apache Spark, Apache Hadoop, Apache Hive에서 Amazon Corretto 11과 17을 모두 지원하지만, 이 버전의 Corretto를 사용하면 일부 워크로드의 성능이 저하될 수 있습니다. 기본값을 변경하기 전에 워크로드를 테스트하는 것이 좋습니다.

다음 표에서는 Amazon EMR 6.15.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
피닉스	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

6.15.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.4.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
delta-standalone-connectors	0.6.0	Delta 커넥터는 Delta Lake를 Flink, Hive, Presto와 같은 엔진과 통합하기 위한 다양한 런타임을 제공합니다.
emr-ddb	5.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.8.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.12.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.29.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.8.0	EMR S3Select 커넥터
emr-wal-cli	1.2.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.60.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.

구성 요소	버전	설명
flink-client	1.17.1-amzn-1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.17.1-amzn-1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.6-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.6-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.6-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.6-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.6-amzn-1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.6-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.3.6-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.6-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.6-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.6-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.6-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.17-amzn-3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-3	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-3	Apache HBase 클러스터용 복구 도구.

구성 요소	버전	설명
hcatalog-client	3.1.3-amzn-8	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-8	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-8	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-8	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-8	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-8	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-8	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.14.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.14.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.14.0-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.14.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.4.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.3	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.283-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.283-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.283-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	426-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	426-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	426-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.4.1-amzn-2	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.4.1-amzn-2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.4.1-amzn-2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.4.1-amzn-2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.08.1-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-6	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-6	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.15.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.15.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services SecondaryNamenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.15.0 변경 로그

6.15.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-11-17	문서 게시	Amazon EMR 6.15.0 릴리스 정보 최초 게시됨
2023-11-17	배포 완료	모든 지원되는 리전에 Amazon EMR 6.15.0 전체 배포됨
2023-11-13	초기 릴리스	최초 상용 리전에 Amazon EMR 6.15.0 최초 배포됨

Amazon EMR 릴리스 6.14.0

6.14.0 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Presto, PigSparkSqoop, Trino, TensorFlowTez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
AWS Java용 SDK	1.12.543	1.12.513	1.12.490	1.12.446
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.4.0	2.2.0
Flink	1.17.1	1.17.0	1.17.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.17	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-2	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.1-amzn-0	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.14.0	emr-6.13.0	emr-6.12.0	emr-6.11.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.3	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.281	0.281	0.281	0.279
Spark	3.4.1	3.4.1	3.4.0	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	422	414	414	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.14.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.14.0에 대한 정보가 포함됩니다. 변경 사항은 6.13.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.14.0 변경 로그 섹션을 참조하세요.

새로운 특성

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- Amazon EMR Managed Scaling은 Amazon EMR 6.14.0 이상에서 생성한 클러스터에 대해 `ap-southeast-3` 아시아 태평양(자카르타) 리전에서 정식 출시되었습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 6.14.0 릴리스는 Amazon EC2에서 실행되는 Amazon EMR을 사용하여 로그 관리를 최적화합니다. 따라서 클러스터 로그의 스토리지 비용이 조금 절감될 수 있습니다.
- 6.14.0 릴리스에서는 Amazon EBS 볼륨에서 크기에 따른 많은 변형을 지원하는 여러 코어 인스턴스를 처리하도록 조정 워크플로를 개선합니다. 이 개선 사항은 코어 노드에만 적용되며, 태스크 노드의 스케일 다운 작업은 영향을 받지 않습니다.
- 6.14.0 릴리스에서는 Amazon EMR이 오픈 소스 애플리케이션(예: Apache Hadoop YARN ResourceManager and HDFS NameNode)과 상호 작용하는 방식을 개선합니다. 이 개선을 통해 클러스터 조정으로 인한 운영 지연 위험이 줄어들고 오픈 소스 애플리케이션과의 연결 문제로 인해 발생하는 시작 장애가 완화됩니다.
- 6.14.0 릴리스에서는 클러스터 시작 시 애플리케이션 설치를 최적화합니다. 이를 통해 특정 조합의 Amazon EMR 애플리케이션에 대한 클러스터 시작 시간이 개선됩니다.
- 6.14.0 릴리스에서는 사용자 지정 도메인이 있는 VPC에서 실행 중인 클러스터에서 코어 또는 태스크 노드가 다시 시작될 때 클러스터 스케일 다운 작업이 중단될 수 있는 문제를 수정합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용]을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부),이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 101.0	4.14.327	2023년 11월 17일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 9월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

6.14.0 기본 Java 버전

Amazon EMR 릴리스 6.12.0 이상에서는 기본적으로 Amazon Corretto 8을 사용하는 모든 애플리케이션을 지원합니다(단, Trino는 예외임). Trino와 관련하여 Amazon EMR은 Amazon EMR 릴리스 6.9.0부터 Amazon Corretto 17을 기본적으로 지원합니다. 또한 Amazon EMR은 Amazon Corretto 11 및 17을 사용하는 일부 애플리케이션도 지원합니다. 이 애플리케이션은 다음 테이블에 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을

사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Amazon EMR은 Apache Spark, Apache Hadoop, Apache Hive에서 Amazon Corretto 11과 17을 모두 지원하지만, 이 버전의 Corretto를 사용하면 일부 워크로드의 성능이 저하될 수 있습니다. 기본값을 변경하기 전에 워크로드를 테스트하는 것이 좋습니다.

다음 표에서는 Amazon EMR 6.14.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	8
PrestoDB	8
Spark	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

6.14.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK

구성 요소	버전	설명
delta	2.4.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
delta-standalone-connectors	0.6.0	Delta 커넥터는 Delta Lake를 Flink, Hive, Presto와 같은 엔진과 통합하기 위한 다양한 런타임을 제공합니다.
emr-ddb	5.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.7.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.11.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.28.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.7.0	EMR S3Select 커넥터
emr-wal-cli	1.1.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.59.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.17.1-amzn-0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
flink-jobmanager-config	1.17.1-amzn-0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-6	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.3.3-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.3-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.17-amzn-2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-2	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-2	Apache HBase 클러스터용 복구 도구.

구성 요소	버전	설명
hcatalog-client	3.1.3-amzn-7	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-7	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-7	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-7	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-7	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-7	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-7	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.13.1-amzn-2	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.13.1-amzn-2	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.13.1-amzn-2	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.13.1-amzn-2	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.3.1-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.3	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.281-amzn-2	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.281-amzn-2	여러 쿼리를 실행하는 서비스.
presto-client	0.281-amzn-2	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	422-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	422-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	422-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.4.1-amzn-1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.4.1-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.4.1-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.4.1-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.06.0-amzn-2	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-5	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-5	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.14.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.14.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.14.0 변경 로그

6.14.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-11-02	배포 완료	모든 지원되는 리전에 Amazon EMR 6.14.0 전체 배포됨
2023-10-10	문서 게시	Amazon EMR 6.14.0 릴리스 정보 최초 게시됨
2023-10-04	초기 릴리스	최초 상용 리전에 Amazon EMR 6.14.0 최초 배포됨

Amazon EMR 릴리스 6.13.0

6.13.0 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
AWS Java용 SDK	1.12.513	1.12.490	1.12.446	1.12.446
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.4.0	2.2.0	2.2.0
Flink	1.17.0	1.17.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.17	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-1	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.11.0
Iceberg	1.3.0-amzn-1	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.13.0	emr-6.12.0	emr-6.11.1	emr-6.11.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.3	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.281	0.281	0.279	0.279
Spark	3.4.1	3.4.0	3.3.2	3.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	414	414	410	410
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.13.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.13.0에 대한 정보가 포함됩니다. 변경 사항은 6.12.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.13.0 변경 로그 섹션을 참조하세요.

새로운 특성

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn 이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값) 로 설정합니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 6.13.0 릴리스에서는 클러스터 종료 명령 실행 시 모든 로그가 Amazon S3에 정기적으로 업로드되도록 Amazon EMR 로그 관리 대몬(daemon)을 개선합니다. 이를 통해 클러스터 종료 속도가 빨라집니다.
- 6.13.0 릴리스에서는 모든 로그 파일을 Amazon S3에 일관되고 시기적절하게 업로드하도록 Amazon EMR 로그 관리 기능을 개선합니다. 이는 특히 장기 실행 EMR 클러스터에 유용합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용](#)을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

6.13.0 기본 Java 버전

Amazon EMR 릴리스 6.12.0 이상에서는 기본적으로 Amazon Corretto 8을 사용하는 모든 애플리케이션을 지원합니다(단, Trino는 예외임). Trino와 관련하여 Amazon EMR은 Amazon EMR 릴리스 6.9.0부터 Amazon Corretto 17을 기본적으로 지원합니다. 또한 Amazon EMR은 Amazon Corretto 11 및 17을 사용하는 일부 애플리케이션도 지원합니다. 이 애플리케이션은 다음 테이블에 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을

[사용하도록 애플리케이션 구성]의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수 있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Amazon EMR은 Apache Spark, Apache Hadoop, Apache Hive에서 Amazon Corretto 11과 17을 모두 지원하지만, 이 버전의 Corretto를 사용하면 일부 워크로드의 성능이 저하될 수 있습니다. 기본값을 변경하기 전에 워크로드를 테스트하는 것이 좋습니다.

다음 표에서는 Amazon EMR 6.13.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	8
PrestoDB	8
Spark	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Spark RAPIDS	17, 11, 8
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

6.13.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK

구성 요소	버전	설명
delta	2.4.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
delta-standalone-connectors	0.6.0	Delta 커넥터는 Delta Lake를 Flink, Hive, Presto와 같은 엔진과 통합하기 위한 다양한 런타임을 제공합니다.
emr-ddb	5.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.6.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.10.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.27.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.6.0	EMR S3Select 커넥터
emr-wal-cli	1.1.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.58.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.17.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
flink-jobmanager-config	1.17.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-5	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.3.3-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.3-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.17-amzn-1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-1	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-1	Apache HBase 클러스터용 복구 도구.

구성 요소	버전	설명
hcatalog-client	3.1.3-amzn-6	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-6	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-6	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-6	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-6	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-6	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-6	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.13.1-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.13.1-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.13.1-amzn-1	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.13.1-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.3.0-amzn-1	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.3	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.281-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.281-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.281-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	414-amzn-1	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	414-amzn-1	여러 쿼리를 실행하는 서비스.
trino-client	414-amzn-1	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.4.1-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.4.1-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.4.1-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.4.1-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.06.0-amzn-1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-4	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-4	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.13.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.13.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.13.0 변경 로그

6.13.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-09-23	배포 완료	모든 지원되는 리전에 Amazon EMR 6.13.0 전체 배포됨
2023-09-12	문서 게시	Amazon EMR 6.13.0 릴리스 정보 최초 게시됨
2023-09-01	초기 릴리스	최초 상용 리전에 Amazon EMR 6.13.0 최초 배포됨

Amazon EMR 릴리스 6.12.0

6.12.0 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Presto, PigSparkSqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
AWS Java용 SDK	1.12.490	1.12.446	1.12.446	1.12.397
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.4.0	2.2.0	2.2.0	2.2.0
Flink	1.17.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.17	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.1-amzn-0	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.11.0	4.10.0
Iceberg	1.3.0-amzn-0	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1

	emr-6.12.0	emr-6.11.1	emr-6.11.0	emr-6.10.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.3	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.281	0.279	0.279	0.278
Spark	3.4.0	3.3.2	3.3.2	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	414	410	410	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.12.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.12.0에 대한 정보가 포함됩니다. 변경 사항은 6.11.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.12.0 변경 로그 섹션을 참조하세요.

새로운 특성

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Amazon EMR 릴리스 6.12.0 이상에서는 Apache Livy, HiveServer2(HS2)를 통한 Apache Hive, Trino, Presto, Hue와의 LDAP 통합을 지원합니다. 6.12.0 이상을 사용하는 EMR 클러스터에 Apache

Spark 및 Apache Hadoop을 설치하고 LDAP를 사용하도록 구성할 수도 있습니다. 자세한 내용은 Amazon EMR에서 인증을 위해 Active Directory 또는 LDAP 서버 사용을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Amazon EMR 릴리스 6.12.0 이상에서는 Flink에 대한 Java 11 런타임 지원을 제공합니다. 자세한 내용은 Java 11과 함께 실행하도록 Flink 구성 단원을 참조하십시오.
- 6.12.0 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항을 통해 크기 조정 작업의 한 번 실패로 인해 클러스터 크기 조정이 무기한 중단될 위험이 줄어듭니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.
- 6.12.0 릴리스는 정상적으로 사용 중지되는 코어 노드가 완전히 해제되기 전에 어떤 이유로든 비정상 상태가 되면 클러스터 스케일 다운 작업이 중단될 수 있는 문제를 수정합니다.
- 6.12.0 릴리스에서는 클러스터가 클러스터의 HDFS 복제 계수 설정 미만으로 코어 노드의 스케일 다운을 시도하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이는 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 6.12.0 릴리스에서는 인스턴스의 상태 변경을 로깅하는 속도를 높여 Amazon EMR에 대한 상태 모니터링 서비스의 성능과 효율성을 개선합니다. 이 개선 사항을 통해 여러 사용자 지정 클라이언트 도구 또는 타사 애플리케이션을 실행하는 클러스터 노드에서 성능이 저하될 가능성이 줄어듭니다.
- 6.12.0 릴리스에서는 Amazon EMR에 대한 클러스터 내 로그 관리 대몬(daemon)의 성능을 개선합니다. 따라서 동시성이 높은 단계를 실행하는 EMR 클러스터를 사용하면 성능이 저하될 가능성이 줄어듭니다.
- Amazon EMR 릴리스 6.12.0에서는 로컬 인스턴스 스토리지의 열린 파일 핸들 및 관련 프로세스에서 현재 사용 중인 모든 로그를 식별하도록 로그 관리 대몬(daemon)이 업그레이드되었습니다. 이 업그

레이드를 통해 Amazon EMR은 로그를 Amazon S3에 아카이브한 후 파일을 올바르게 삭제하고 스토리지 공간을 확보할 수 있습니다.

- 6.12.0 릴리스에는 로컬 클러스터 파일 시스템에서 사용되지 않는 빈 단계 디렉터리를 삭제하는 로그 관리 대몬(daemon) 개선 사항이 포함되어 있습니다. 빈 디렉터리가 너무 많으면 Amazon EMR 대몬(daemon)의 성능이 저하되고 디스크가 과도하게 사용될 수 있습니다.
- 6.12.0 릴리스에서는 YARN 타임라인 서버 로그의 로그 로테이션을 지원합니다. 이를 통해 특히 장기 실행 클러스터의 경우 디스크 과다 사용 시나리오를 최소화할 수 있습니다.
- Amazon EMR 6.10.0 이상에서는 기본 루트 볼륨 크기가 15GB로 늘어났습니다. 이전 릴리스에서 기본 루트 볼륨 크기는 10GB였습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

6.12.0 기본 Java 버전

Amazon EMR 릴리스 6.12.0 이상에서는 기본적으로 Amazon Corretto 8을 사용하는 모든 애플리케이션을 지원합니다(단, Trino는 예외임). Trino와 관련하여 Amazon EMR은 Amazon EMR 릴리스 6.9.0부터 Amazon Corretto 17을 기본적으로 지원합니다. 또한 Amazon EMR은 Amazon Corretto 11 및 17을 사용하는 일부 애플리케이션도 지원합니다. 이 애플리케이션은 다음 테이블에 나와 있습니다. 클러스터의 기본 JVM을 변경하려면 클러스터에서 실행되는 각 애플리케이션에 대해 특정 Java 가상 머신을 사용하도록 애플리케이션 구성의 지침을 따릅니다. 클러스터에서는 Java 런타임 버전 1개만 사용할 수

있습니다. Amazon EMR은 동일한 클러스터에서 다양한 런타임 버전을 사용해 서로 다른 노드 또는 애플리케이션을 실행하는 것을 지원하지 않습니다.

Amazon EMR은 Apache Spark, Apache Hadoop, Apache Hive에서 Amazon Corretto 11과 17을 모두 지원하지만, 이 버전의 Corretto를 사용하면 일부 워크로드의 성능이 저하될 수 있습니다. 기본값을 변경하기 전에 워크로드를 테스트하는 것이 좋습니다.

다음 표에서는 Amazon EMR 6.12.0에 있는 애플리케이션의 기본 Java 버전을 확인할 수 있습니다.

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
델타	17, 11, 8
Flink	11, 8
Ganglia	8
HBase	11, 8
HCatalog	17, 11, 8
Hadoop	17, 11, 8
Hive	17, 11, 8
Hudi	17, 11, 8
Iceberg	17, 11, 8
Livy	17, 11, 8
Oozie	17, 11, 8
피닉스	8
PrestoDB	8
Spark	17, 11, 8
Spark RAPIDS	17, 11, 8

Application	Java 및 Amazon Corretto 버전(기본값은 굵은체로 표시)
Sqoop	8
Tez의 Hive 실행 시간 비교	17, 11, 8
Trino	17
Zeppelin	8
Pig	8
Zookeeper	8

6.12.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.4.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
delta-standalone-connectors	0.6.0	Delta 커넥터는 Delta Lake를 Flink, Hive, Presto와 같은 엔진과 통합하기 위한 다양한 런타임을 제공합니다.
emr-ddb	5.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.5.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.9.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.26.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.5.0	EMR S3Select 커넥터
emr-wal-cli	1.1.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.57.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.17.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.17.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.3-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	3.3.3-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.17-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.17-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.17-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.4.17-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.17-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.17-amzn-0	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-5	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	3.1.3-amzn-5	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-5	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-5	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-5	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-5	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-5	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.13.1-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.13.1-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.13.1-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.13.1-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
iceberg	1.3.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.7.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.3	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.281-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.281-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.281-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	414-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	414-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	414-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.4.0-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.4.0-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.4.0-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.4.0-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.06.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-3	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-3	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.12.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.12.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-java-home	Hadoop의 KMS java 홈 변경	Not available.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.12.0 변경 로그

6.12.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-07-27	설명서 업데이트	6.12용 Java 옵션을 업데이트하고 Oozie 튜토리얼을 추가하여 JVM 업데이트하십시오.
2023-07-21	배포 완료	모든 지원되는 리전에 Amazon EMR 6.12.0 전체 배포됨
2023-07-21	문서 게시	Amazon EMR 6.12.0 릴리스 정보 최초 게시됨
2023-07-12	초기 릴리스	최초 상용 리전에 Amazon EMR 6.12.0 최초 배포됨

Amazon EMR 릴리스 6.11.1

6.11.1 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, HudiIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전

- [Amazon EMR 6.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
AWS Java용 SDK	1.12.446	1.12.446	1.12.397	1.12.397
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.2.0
Flink	1.16.0	1.16.0	1.16.0	1.16.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.15
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0
Hue	4.11.0	4.11.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0

	emr-6.11.1	emr-6.11.0	emr-6.10.1	emr-6.10.0
JupyterHub	1.4.1	1.4.1	1.5.0	1.5.0
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.279	0.279	0.278	0.278
Spark	3.3.2	3.3.2	3.3.1	3.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.11.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	410	410	403	403
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.11.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.11.1에 대한 정보가 포함됩니다. 변경 사항은 6.11.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.11.1 변경 로그 섹션을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 잠금 경합으로 인해 노드가 서비스 해제를 시도하면서 동시에 노드가 추가되거나 제거되면 노드가 교착 상태에 빠질 수 있습니다. 따라서 Hadoop Resource Manager(YARN)가 응답하지 않게 되고 수신 및 현재 실행 중인 모든 컨테이너에 영향을 미칩니다.
- 이 릴리스에는 다시 시작한 후 고가용성 클러스터를 장애 상태에서 복구할 수 있도록 지원하는 변경 사항이 포함되어 있습니다.
- 이 릴리스에는 Hue 및 HBase에 대한 보안 수정 사항이 포함되어 있습니다.
- 이 릴리스에서는 Amazon EMR을 사용하여 Spark에서 워크로드를 실행하는 클러스터가 `contains`, `startsWith`, `endsWith`, `like`에서 잘못된 결과를 자동으로 수신할 수 있는 문제를 수정했습니다. 이 문제는 Amazon EMR Hive3 Metastore Server(HMS)에 메타데이터가 있는 파티셔닝된 필드에서 표현식을 사용할 때 발생합니다.
- 이 릴리스에서는 사용자 정의 함수(UDF)가 없을 때 발생하는 Glue 측 제한 관련 문제를 수정했습니다.
- 이 릴리스에서는 YARN 서비스 해제 시 로그 푸셔가 컨테이너 로그를 S3로 푸시하기 전에 노드 로그 집계 서비스에서 컨테이너 로그를 삭제하는 문제를 수정했습니다.
- 이 릴리스에서는 Hadoop에 노드 레이블이 활성화된 경우 FairShare 스케줄러 지표와 관련된 문제를 수정했습니다.
- 이 릴리스에서는 `spark-defaults.conf`에서 `spark.yarn.heterogeneousExecutors.enabled` 구성의 기본 `true` 값을 설정할 때 Spark 성능에 영향을 주던 문제를 수정했습니다.
- 이 릴리스에서는 Reduce 작업이 셔플 데이터를 읽지 못하는 문제를 수정했습니다. 이 문제로 인해 Hive 쿼리가 실패하고 메모리 손상 오류가 발생했습니다.
- 이 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항을 통해 크기 조정 작업의 한 번 실패로 인해 클러스터

크기 조정이 무기한 중단될 위험이 줄어듭니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.

- 이 릴리스에서는 클러스터가 클러스터의 HDFS 복제 계수 설정 미만으로 코어 노드의 스케일 다운을 시도하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이는 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 로컬 인스턴스 스토리지의 열린 파일 핸들 및 관련 프로세스에서 현재 사용 중인 모든 로그를 식별하도록 로그 관리 대몬(daemon)이 업그레이드되었습니다. 이 업그레이드를 통해 Amazon EMR은 로그를 Amazon S3에 아카이브한 후 파일을 올바르게 삭제하고 스토리지 공간을 확보할 수 있습니다.
- 이 릴리스에는 로컬 클러스터 파일 시스템에서 사용되지 않는 빈 단계 디렉터리를 삭제하는 로그 관리 대몬(daemon) 개선 사항이 포함되어 있습니다. 빈 디렉터리가 너무 많으면 Amazon EMR 대몬(daemon)의 성능이 저하되고 디스크가 과도하게 사용될 수 있습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용]을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaLabel(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

6.11.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.2.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
delta-standalone-connectors	0.6.0	Delta 커넥터는 Delta Lake를 Flink, Hive, Presto와 같은 엔진과 통합하기 위한 다양한 런타임을 제공합니다.
emr-ddb	5.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.4.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.8.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.25.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.4.0	EMR S3Select 커넥터
emr-wal-cli	1.1.0	emrwal 목록/삭제에 사용되는 CLI입니다.

구성 요소	버전	설명
emrfs	2.56.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.16.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.16.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-3.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-3.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-3.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-3.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-3.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	3.3.3-amzn-3.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.3-amzn-3.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.3-amzn-3.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-3.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-3.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-3.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.15-amzn-1.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.15-amzn-1.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.15-amzn-1.1	HBase 명령줄 클라이언트
hbase-rest-server	2.4.15-amzn-1.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.15-amzn-1.1	HBase에 Thrift 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hbase-operator-tools	2.4.15-amzn-1.1	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-4.1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-4.1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-4.1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-4.1	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-4.1	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-4.1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-4.1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.13.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.13.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.13.0-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.13.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.2.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.2	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.279-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.279-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.279-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	410-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	410-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	410-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.3.2-amzn-0.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.3.2-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.2-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.2-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.02.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-2.1	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-2.1	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.11.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.11.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.11.1 변경 로그

6.11.1 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-30	릴리스 정보 업데이트	릴리스 정보에 컨트롤 플레인 관련 수정 사항이 여러 개 추가되었습니다.
2023-08-21	문서 게시	Amazon EMR 6.11.1 릴리스 정보 최초 게시됨
2023-08-16	배포 완료	모든 지원되는 리전에 Amazon EMR 6.11.1 전체 배포됨
2023-08-04	초기 릴리스	제한된 상용 리전에 Amazon EMR 6.11.1 최초 배포됨

Amazon EMR 릴리스 6.11.0

6.11.0 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, Trino및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전

- [Amazon EMR 6.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
AWS Java용 SDK	1.12.446	1.12.397	1.12.397	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.2.0	2.1.0
Flink	1.16.0	1.16.0	1.16.0	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.15	2.4.13
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.13.0-amzn-0	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0
Hue	4.11.0	4.10.0	4.10.0	4.10.0
Iceberg	1.2.0-amzn-0	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0

	emr-6.11.0	emr-6.10.1	emr-6.10.0	emr-6.9.1
JupyterHub	1.4.1	1.5.0	1.5.0	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.279	0.278	0.278	0.276
Spark	3.3.2	3.3.1	3.3.1	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.11.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	410	403	403	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.11.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.11.0에 대한 정보가 포함됩니다. 변경 사항은 6.10.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

새로운 특성

- Amazon EMR 6.11.0에서는 Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0, PrestoDB 0.279-amzn-0을 지원합니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 TRUE (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Amazon EMR 6.11.0에서 DynamoDB 커넥터가 버전 5.0.0으로 업그레이드되었습니다. 버전 5.0.0은를 사용합니다 AWS SDK for Java 2.x. 이전 릴리스는 AWS SDK for Java 1.x를 사용했습니다. 이번 업그레이드로, Amazon EMR 6.11에서 DynamoDB 커넥터를 사용하기 전에 코드를 테스트해 보는 것이 좋습니다.
- Amazon EMR 6.11.0용 DynamoDB 커넥터가 DynamoDB 서비스를 직접 호출할 때는 사용자가 `dynamodb.endpoint` 속성에 제공한 리전 값을 사용합니다. `dynamodb.endpoint` 사용 시 `dynamodb.region`도 구성하고, 두 속성이 모두 동일한 AWS 리전을 대상으로 지정하는 것이 좋습니다. `dynamodb.endpoint`를 사용하고 `dynamodb.region`를 구성하지 않으면 Amazon EMR 6.11.0용 DynamoDB 커넥터가 잘못된 리전 예외를 반환하고 Amazon EC2 인스턴스 메타데이터 서비스(IMDS)에서 AWS 리전 정보를 조정하려고 시도합니다. 커넥터가 IMDS에서 리전을 검색할 수 없는 경우 기본값은 미국 동부(버지니아 북부)(`us-east-1`)입니다. 다음 오류는 `dynamodb.region` 속성을 올바르게 구성하지 않을 경우 발생할 수 있는 잘못된 리전 예외의 예입니다. 2.x로 AWS SDK for Java 업그레이드의 영향을 받는 클래스에 대한 `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` 자세한 내용은 Amazon EMR - DynamoDB 커넥터의 GitHub 리포지토리에서 1.x AWS SDK for Java 에서 2.x(#175)로 업그레이드 커밋을 참조하세요.

- 이 릴리스에서는 열 이름 바꾸기 작업 후 Delta Lake를 사용하여 Amazon S3에 델타 테이블 데이터를 저장할 때 열 데이터가 NULL로 표시되는 문제를 수정했습니다. Delta Lake의 이 실험 기능에 대한 자세한 내용은 Delta Lake 사용 설명서에서 Column rename operation을 참조하세요.
- 6.11.0 릴리스에서는 여러 프라이머리 노드를 포함하는 클러스터에서 프라이머리 노드 중 하나를 복제하여 엣지 노드를 생성할 때 발생할 수 있는 문제를 수정했습니다. 복제된 엣지 노드로 인해 스케일 다운 작업이 지연되거나 프라이머리 노드의 메모리 사용률이 높아질 수 있습니다. EMR 클러스터와 통신하기 위해 엣지 노드를 생성하는 방법에 대한 자세한 내용은 GitHub의 `aws-samples` 리포지토리에서 Edge Node Creator를 참조하세요.
- 6.11.0 릴리스에서는 Amazon EMR이 재부팅 후 Amazon EBS 볼륨을 인스턴스에 다시 마운트하는 데 사용하는 자동화 프로세스를 개선합니다.
- 6.11.0 릴리스에서는 Amazon EMR이 Amazon CloudWatch에 게시하는 Hadoop 지표에 간헐적으로 간격이 발생하는 문제를 수정했습니다.
- 6.11.0 릴리스에서는 클러스터의 노드 제외 목록이 포함된 YARN 구성 파일에 대한 업데이트가 디스크 과다 사용으로 인해 중단되는 EMR 클러스터 관련 문제를 수정했습니다. 업데이트가 불완전하면 향후 클러스터 스케일 다운 작업에 방해가 됩니다. 이 릴리스에서는 클러스터를 정상 상태로 유지하고 조정 작업을 예상대로 수행하도록 보장합니다.
- Amazon EMR 6.10.0 이상에서는 기본 루트 볼륨 크기가 15GB로 늘어났습니다. 이전 릴리스에서 기본 루트 볼륨 크기는 10GB였습니다.
- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` 값이 `false`로 설정됩니다.

  이 수정 사항은 YARN-9608 때문에 발생한 문제를 해결하지만, Managed Scaling이 활성화된 클러스터에서는 셔플 데이터 손실로 인해 Hive 작업이 실패할 수 있습니다. 이번 릴리스에서는 Hive 워크로드에 대해 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data`도 설정하여 이러한 위험을 완화했습니다. 이 구성은 Amazon EMR 릴리스 6.11.0 이상에서만 사용할 수 있습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

Note

이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, `ListReleaseLabels` API 또는 `list-release-labels` CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

6.11.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.2.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
delta-standalone-connectors	0.6.0	Delta 커넥터는 Delta Lake를 Flink, Hive, Presto와 같은 엔진과 통합하기 위한 다양한 런타임을 제공합니다.
emr-ddb	5.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.4.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.8.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.25.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.4.0	EMR S3Select 커넥터

구성 요소	버전	설명
emr-wal-cli	1.1.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.56.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.16.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.16.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-hdfs-journalnode	3.3.3-amzn-3	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.3-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.3-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.15-amzn-1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.15-amzn-1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.15-amzn-1	HBase 명령줄 클라이언트
hbase-rest-server	2.4.15-amzn-1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hbase-thrift-server	2.4.15-amzn-1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.15-amzn-1	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-4	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-4	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.13.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.13.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.13.0-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hudi-spark	0.13.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.2.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-connectors	5.1.2	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.279-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.279-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.279-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	410-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	410-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	410-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템

구성 요소	버전	설명
spark-client	3.3.2-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.3.2-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.2-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.2-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	23.02.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-2	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-2	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.11.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.11.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services SecondaryNamenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.11.0 변경 로그

6.11.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-21	업데이트	Hadoop 3.3.3에서 발생하는 문제가 해결되었습니다.
2023-07-26	업데이트	새 OS 릴리스 레이블 `2.0.20230612.0` 및 `2.0.20230628.0` .
2023-06-09	배포 완료	모든 지원되는 리전에 Amazon EMR 6.11.0 전체 배포됨
2023-06-09	문서 게시	Amazon EMR 6.11.0 릴리스 정보 최초 게시됨
2023-06-08	초기 릴리스	최초 상용 리전에 Amazon EMR 6.11.0 최초 배포됨

Amazon EMR 릴리스 6.10.1

6.10.1 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hive, Hadoop, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
AWS Java용 SDK	1.12.397	1.12.397	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.2.0	2.1.0	2.1.0
Flink	1.16.0	1.16.0	1.15.2	1.15.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.15	2.4.13	2.4.13
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.3.3
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0

	emr-6.10.1	emr-6.10.0	emr-6.9.1	emr-6.9.0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.6.0
JupyterHub	1.5.0	1.5.0	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.278	0.278	0.276	0.276
Spark	3.3.1	3.3.1	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.11.0	2.10.0	2.10.0
Tez	0.10.2	0.10.2	0.10.2	0.10.2
Trino (PrestoSQL)	403	403	398	398
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.10.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.10.1에 대한 정보가 포함됩니다. 변경 사항은 6.10.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.10.1 변경 로그 섹션을 참조하세요.

알려진 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태ExecutorState라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- 잠금 경합으로 인해 노드가 서비스 해제를 시도하면서 동시에 노드가 추가되거나 제거되면 노드가 교착 상태에 빠질 수 있습니다. 따라서 Hadoop Resource Manager(YARN)가 응답하지 않게 되고 수신 및 현재 실행 중인 모든 컨테이너에 영향을 미칩니다.
- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` 값이 `false`로 설정됩니다.

  이 수정 사항은 YARN-9608 때문에 발생한 문제를 해결하지만, Managed Scaling이 활성화된 클러스터에서는 셔플 데이터 손실로 인해 Hive 작업이 실패할 수 있습니다. 이번 릴리스에서는 Hive 워크로드에 대해 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data`도 설정하여 이러한 위험을 완화했습니다. 이 구성은 Amazon EMR 릴리스 6.11.0 이상에서만 사용할 수 있습니다.
- 지표 수집기는 인스턴스 그룹 구성을 사용하는 클러스터의 프라이머리 노드를 장애 조치한 후 컨트롤 플레인에 지표를 전송하지 않습니다.
- 이 릴리스에는 다시 시작한 후 고가용성 클러스터를 장애 상태에서 복구할 수 있도록 지원하는 변경 사항이 포함되어 있습니다.
- 이 릴리스에는 Hue 및 HBase에 대한 보안 수정 사항이 포함되어 있습니다.

- 이 릴리스에서는 Amazon EMR을 사용하여 Spark에서 워크로드를 실행하는 클러스터가 `contains`, `startsWith`, `endsWith`, `like`에서 잘못된 결과를 자동으로 수신할 수 있는 문제를 수정했습니다. 이 문제는 Amazon EMR Hive3 Metastore Server(HMS)에 메타데이터가 있는 파티셔닝된 필드에서 표현식을 사용할 때 발생합니다.
- 이 릴리스에서는 사용자 정의 함수(UDF)가 없을 때 발생하는 Glue 측 제한 관련 문제를 수정했습니다.
- 이 릴리스에서는 YARN 서비스 해제 시 로그 푸셔가 컨테이너 로그를 S3로 푸시하기 전에 노드 로그 집계 서비스에서 컨테이너 로그를 삭제하는 문제를 수정했습니다.
- 이 릴리스에서는 Hadoop에 노드 레이블이 활성화된 경우 FairShare 스케줄러 지표와 관련된 문제를 수정했습니다.
- 이 릴리스에서는 `spark-defaults.conf`에서 `spark.yarn.heterogeneousExecutors.enabled` 구성의 기본 `true` 값을 설정할 때 Spark 성능에 영향을 주던 문제를 수정했습니다.
- 이 릴리스에서는 Reduce 작업이 셔플 데이터를 읽지 못하는 문제를 수정했습니다. 이 문제로 인해 Hive 쿼리가 실패하고 메모리 손상 오류가 발생했습니다.
- 이 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항을 통해 크기 조정 작업의 한 번 실패로 인해 클러스터 크기 조정이 무기한 중단될 위험이 줄어듭니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.
- 이 릴리스에서는 클러스터가 클러스터의 HDFS 복제 계수 설정 미만으로 코어 노드의 스케일 다운을 시도하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이는 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 로컬 인스턴스 스토리지의 열린 파일 핸들 및 관련 프로세스에서 현재 사용 중인 모든 로그를 식별하도록 로그 관리 대몬(daemon)이 업그레이드되었습니다. 이 업그레이드를 통해 Amazon EMR은 로그를 Amazon S3에 아카이브한 후 파일을 올바르게 삭제하고 스토리지 공간을 확보할 수 있습니다.
- 이 릴리스에는 로컬 클러스터 파일 시스템에서 사용되지 않는 빈 단계 디렉터리를 삭제하는 로그 관리 대몬(daemon) 개선 사항이 포함되어 있습니다. 빈 디렉터리가 너무 많으면 Amazon EMR 대몬(daemon)의 성능이 저하되고 디스크가 과도하게 사용될 수 있습니다.
- 이 릴리스에서는 여러 프라이머리 노드를 포함하는 클러스터에서 프라이머리 노드 중 하나를 복제하여 엣지 노드를 생성할 때 발생할 수 있는 문제를 수정했습니다. 복제된 엣지 노드로 인해 스케일 다운 작업이 지연되거나 프라이머리 노드의 메모리 사용률이 높아질 수 있습니다. EMR 클러스터와 통신하기 위해 엣지 노드를 생성하는 방법에 대한 자세한 내용은 GitHub의 `aws-samples` 리포지토리에서 Edge Node Creator를 참조하세요.

- 이 릴리스에서는 Amazon EMR이 재부팅 후 Amazon EBS 볼륨을 인스턴스에 다시 마운트하는 데 사용하는 자동화 프로세스를 개선합니다.
- 이 릴리스에서는 Amazon EMR이 Amazon CloudWatch에 게시하는 Hadoop 지표에 간헐적으로 간격이 발생하는 문제를 수정했습니다.
- 이 릴리스에서는 클러스터의 노드 제외 목록이 포함된 YARN 구성 파일에 대한 업데이트가 디스크 과다 사용으로 인해 중단되는 EMR 클러스터 관련 문제를 수정했습니다. 업데이트가 불완전하면 향후 클러스터 스케일 다운 작업에 방해가 됩니다. 이 릴리스에서는 클러스터를 정상 상태로 유지하고 조정 작업을 예상대로 수행하도록 보장합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용](#)을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

6.10.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.2.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.7.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.24.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.3.0	EMR S3Select 커넥터
emr-wal-cli	1.0.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.55.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.16.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
flink-jobmanager-config	1.16.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-2.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-2.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-2.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-2.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-2.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-2.1	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.3.3-amzn-2.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.3-amzn-2.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-2.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-2.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-2.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.15-amzn-0.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.15-amzn-0.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.15-amzn-0.1	HBase 명령줄 클라이언트
hbase-rest-server	2.4.15-amzn-0.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.15-amzn-0.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.15-amzn-0.1	Apache HBase 클러스터용 복구 도구.

구성 요소	버전	설명
hcatalog-client	3.1.3-amzn-3.1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-3.1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-3.1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-3.1	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-3.1	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-3.1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-3.1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.12.2-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.12.2-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.12.2-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.12.2-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.1.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.278.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.278.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.278.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	403-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	403-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	403-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.3.1-amzn-0.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.3.1-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.1-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.1-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.12.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-1.1	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-1.1	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.10.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.10.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.10.1 변경 로그

6.10.1 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-30	릴리스 정보 업데이트	릴리스 정보에 컨트롤 플레인 관련 수정 사항이 여러 개 추가되었습니다.
2023-08-21	문서 게시	Amazon EMR 6.10.1 릴리스 정보 최초 게시됨
2023-08-16	배포 완료	모든 지원되는 리전에 Amazon EMR 6.10.1 전체 배포됨
2023-08-04	초기 릴리스	제한된 상용 리전에 Amazon EMR 6.10.1 최초 배포됨

Amazon EMR 릴리스 6.10.0

6.10.0 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Pig, Presto, Spark SqoopTrino, TensorFlowTez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전

- [Amazon EMR 6.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
AWS Java용 SDK	1.12.397	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.2.0	2.1.0	2.1.0	-
Flink	1.16.0	1.15.2	1.15.2	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.15	2.4.13	2.4.13	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.3.3	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.2-amzn-0	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	1.1.0-amzn-0	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.6.0	2.1.0

	emr-6.10.0	emr-6.9.1	emr-6.9.0	emr-6.8.1
JupyterHub	1.5.0	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.278	0.276	0.276	0.273
Spark	3.3.1	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.11.0	2.10.0	2.10.0	2.9.1
Tez	0.10.2	0.10.2	0.10.2	0.9.2
Trino (PrestoSQL)	403	398	398	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.10.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.10.0에 대한 정보가 포함됩니다. 변경 사항은 6.9.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

새로운 특성

- Amazon EMR 6.10.0에서는 Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403, PrestoDB 0.278.1을 지원합니다.
- Amazon EMR 6.10.0에는 Hudi 테이블의 데이터에 대한 읽기 액세스를 제공하는 네이티브 Trino-Hudi 커넥터가 포함되어 있습니다. `trino-cli --catalog hudi`를 사용하여 커넥터를 활성화하고, `trino-connector-hudi`를 사용하여 요구 사항에 맞게 커넥터를 구성할 수 있습니다. Amazon EMR과의 기본 통합을 통해 더 이상 Hudi 테이블을 쿼리하는 데 `trino-connector-hive`를 사용하지 않아도 됩니다. 새 커넥터에서 지원되는 구성 목록은 Trino 설명서에서 Hudi connector 페이지를 참조하세요.
- Amazon EMR 릴리스 6.10.0 이상에서는 Apache Flink와의 Apache Zeppelin 통합을 지원합니다. 자세한 정보는 Amazon EMR의 Zeppelin에서 Flink 작업을 참조하세요.

알려진 문제

- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 6.10.0에서 이 문제를 해결하려면 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications`의 값을 `false`로 설정하면 됩니다. Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 기본적으로 구성이 `false`로 설정되어 있습니다.
- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정합니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 6.10.0은 `minimal-json.jar`에서 Apache Spark용 Amazon Redshift 통합의 종속 항목을 제거하고 필요한 Spark-Redshift 관련 jar(`spark-redshift.jar`, `spark-avro.jar`, `RedshiftJDBC.jar`)을 Spark의 실행기 클래스 경로에 자동으로 추가합니다.
- 6.10.0 릴리스에서는 MR 클러스터의 추가 로그 폴더를 모니터링하도록 클러스터 내 로그 관리 대몬(daemon)이 개선되었습니다. 이 개선 사항은 디스크 과다 사용 시나리오를 최소화합니다.
- 6.10.0 릴리스에서는 클러스터 내 로그 관리 대몬(daemon)이 중지되면 해당 대몬(daemon)을 자동으로 다시 시작합니다. 이 개선 사항은 디스크 과다 사용으로 인해 노드가 비정상으로 보이는 위험을 줄입니다.
- Amazon EMR 6.10.0은 EMRFS 사용자 매핑에 대한 리전 엔드포인트를 지원합니다.
- Amazon EMR 6.10.0 이상에서는 기본 루트 볼륨 크기가 15GB로 늘어났습니다. 이전 릴리스에서 기본 루트 볼륨 크기는 10GB였습니다.
- 6.10.0 릴리스에서는 다른 모든 Spark 실행기가 YARN 리소스 관리자의 서비스 해제 호스트에 있을 때 해당 Spark 작업이 중단되는 문제를 수정했습니다.
- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 HIVE-20703 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 HIVE-22269 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

> **Note**
>
> 이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, `ListReleaseLabels` API 또는 `list-release-labels` CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)
2.0.2023 404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

6.10.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.2.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.7.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.24.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.3.0	EMR S3Select 커넥터
emr-wal-cli	1.0.0	emrwal 목록/삭제에 사용되는 CLI입니다.
emrfs	2.55.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.16.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
flink-jobmanager-config	1.16.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-2	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.3.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.3.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.15-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.15-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.15-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.4.15-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.15-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.15-amzn-0	Apache HBase 클러스터용 복구 도구.

구성 요소	버전	설명
hcatalog-client	3.1.3-amzn-3	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-3	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-3	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-3	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-3	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-3	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-3	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.12.2-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.12.2-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.12.2-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.12.2-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	1.1.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.5.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.8.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.278.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.278.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.278.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	403-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	403-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	403-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.3.1-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.3.1-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.1-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.1-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.12.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-1	tez YARN 애플리케이션 및 라이브러리.
tez-on-worker	0.10.2-amzn-1	워커 노드에 대한 tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.10.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.10.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Ad

분류	설명	재구성 작업
		ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-hudi	Trino의 hudi.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-logback	Ranger KMS의 kms-logback.xml 파일에서 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.10.0 변경 로그

6.10.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-21	업데이트	Hadoop 3.3.3에서 발생하는 것으로 알려진 문제가 추가되었습니다.
2023-07-26	업데이트	새 OS 릴리스 레이블 `2.0.20230612.0` 및 `2.0.20230628.0` .
2023-03-02	배포 완료	모든 지원되는 리전에 Amazon EMR 6.10 전체 배포됨
2023-03-02	문서 게시	Amazon EMR 6.10 릴리스 정보 최초 게시됨
2023-02-27	초기 릴리스	제한된 상용 리전에 Amazon EMR 6.10 최초 배포됨

Amazon EMR 릴리스 6.9.1

6.9.1 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
AWS Java용 SDK	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	2.1.0	-	-
Flink	1.15.2	1.15.2	1.15.1	1.15.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.13	2.4.12	2.4.12
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.3.3	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0

	emr-6.9.1	emr-6.9.0	emr-6.8.1	emr-6.8.0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.9.1
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.276	0.276	0.273	0.273
Spark	3.3.0	3.3.0	3.3.0	3.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.10.0	2.9.1	2.9.1
Tez	0.10.2	0.10.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	398	388	388
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.1
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.10

6.9.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.9.1에 대한 정보가 포함됩니다. 변경 사항은 6.9.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.9.1 변경 로그 섹션을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` 값이 `false`로 설정됩니다.

  이 수정 사항은 YARN-9608 때문에 발생한 문제를 해결하지만, Managed Scaling이 활성화된 클러스터에서는 셔플 데이터 손실로 인해 Hive 작업이 실패할 수 있습니다. 이번 릴리스에서는 Hive 워크로드에 대해 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data`도 설정하여 이러한 위험을 완화했습니다. 이 구성은 Amazon EMR 릴리스 6.11.0 이상에서만 사용할 수 있습니다.
- 지표 수집기는 인스턴스 그룹 구성을 사용하는 클러스터의 프라이머리 노드를 장애 조치한 후 컨트롤 플레인에 지표를 전송하지 않습니다.
- 이 릴리스에서는 지표 수집기 엔드포인트에 대한 실패한 HTTP 요청의 재시도가 제거되었습니다.
- 이 릴리스에는 다시 시작한 후 고가용성 클러스터를 장애 상태에서 복구할 수 있도록 지원하는 변경 사항이 포함되어 있습니다.
- 이 릴리스에서는 사용자가 생성한 UID가 클 경우 오버플로 예외가 발생하는 문제를 수정했습니다.
- 이 릴리스에서는 Amazon EMR 재구성 프로세스와 관련된 제한 시간 초과 문제를 수정했습니다.
- 이 릴리스에는 보안 수정 사항이 포함되어 있습니다.
- 이 릴리스에서는 Amazon EMR을 사용하여 Spark에서 워크로드를 실행하는 클러스터가 `contains`, `startsWith`, `endsWith`, `like`에서 잘못된 결과를 자동으로 수신할 수 있는 문제를 수정했습니다. 이 문제는 Amazon EMR Hive3 Metastore Server(HMS)에 메타데이터가 있는 파티셔닝된 필드에서 표현식을 사용할 때 발생합니다.
- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 HIVE-20703 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 HIVE-22269 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.

- HDFS를 스크래치 디렉터리로 사용하고 작은 파일 병합을 활성화하며 테이블에 정적 파티션 경로가 포함된 Hive에서 데이터 손실이 발생할 수 있습니다.
- 이 릴리스에서는 ETL 작업 종료 시 작은 작은 파일 병합(기본적으로 비활성화됨)이 활성화된 경우 Hive의 성능 문제를 수정했습니다.
- 이 릴리스에서는 사용자 정의 함수(UDF)가 없을 때 발생하는 Glue 측 제한 관련 문제를 수정했습니다.
- 이 릴리스에서는 YARN 서비스 해제 시 로그 푸셔가 컨테이너 로그를 S3로 푸시하기 전에 노드 로그 집계 서비스에서 컨테이너 로그를 삭제하는 문제를 수정했습니다.
- 이 릴리스에서는 HBase에 대한 지속적 스토어 파일 추적에서 압축된 파일 및 아카이브된 파일의 처리를 수정합니다.
- 이 릴리스에서는 `spark-defaults.conf`에서 `spark.yarn.heterogeneousExecutors.enabled` 구성의 기본 `true` 값을 설정할 때 Spark 성능에 영향을 주던 문제를 수정했습니다.
- 이 릴리스에서는 Reduce 작업이 셔플 데이터를 읽지 못하는 문제를 수정했습니다. 이 문제로 인해 Hive 쿼리가 실패하고 메모리 손상 오류가 발생했습니다.
- 이 릴리스에서는 노드 대체 중에 HDFS Namenode(NN) 서비스가 안전 모드에서 멈춘 경우 노드 프로비저너에 오류가 발생하는 문제를 수정했습니다.
- 이 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항을 통해 크기 조정 작업의 한 번 실패로 인해 클러스터 크기 조정이 무기한 중단될 위험이 줄어듭니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.
- 이 릴리스에서는 클러스터가 클러스터의 HDFS 복제 계수 설정 미만으로 코어 노드의 스케일 다운을 시도하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이는 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 로컬 인스턴스 스토리지의 열린 파일 핸들 및 관련 프로세스에서 현재 사용 중인 모든 로그를 식별하도록 로그 관리 대몬(daemon)이 업그레이드되었습니다. 이 업그레이드를 통해 Amazon EMR은 로그를 Amazon S3에 아카이브한 후 파일을 올바르게 삭제하고 스토리지 공간을 확보할 수 있습니다.
- 이 릴리스에는 로컬 클러스터 파일 시스템에서 사용되지 않는 빈 단계 디렉터리를 삭제하는 로그 관리 대몬(daemon) 개선 사항이 포함되어 있습니다. 빈 디렉터리가 너무 많으면 Amazon EMR 대몬(daemon)의 성능이 저하되고 디스크가 과도하게 사용될 수 있습니다.
- 이 릴리스에서는 여러 프라이머리 노드를 포함하는 클러스터에서 프라이머리 노드 중 하나를 복제하여 엣지 노드를 생성할 때 발생할 수 있는 문제를 수정했습니다. 복제된 엣지 노드로 인해 스케일 다운 작업이 지연되거나 프라이머리 노드의 메모리 사용률이 높아질 수 있습니다. EMR 클러스터와

통신하기 위해 엣지 노드를 생성하는 방법에 대한 자세한 내용은 GitHub의 `aws-samples` 리포지토리에서 [Edge Node Creator](https://github.com/aws-samples)를 참조하세요.

- 이 릴리스에서는 Amazon EMR이 재부팅 후 Amazon EBS 볼륨을 인스턴스에 다시 마운트하는 데 사용하는 자동화 프로세스를 개선합니다.
- 이 릴리스에서는 Amazon EMR이 Amazon CloudWatch에 게시하는 Hadoop 지표에 간헐적으로 간격이 발생하는 문제를 수정했습니다.
- 이 릴리스에서는 클러스터의 노드 제외 목록이 포함된 YARN 구성 파일에 대한 업데이트가 디스크 과다 사용으로 인해 중단되는 EMR 클러스터 관련 문제를 수정했습니다. 업데이트가 불완전하면 향후 클러스터 스케일 다운 작업에 방해가 됩니다. 이 릴리스에서는 클러스터를 정상 상태로 유지하고 조정 작업을 예상대로 수행하도록 보장합니다.
- 이 릴리스에서는 MR 클러스터의 추가 로그 폴더를 모니터링하도록 클러스터 내 로그 관리 대몬(daemon)이 개선되었습니다. 이 개선 사항은 디스크 과다 사용 시나리오를 최소화합니다.
- 이 릴리스에서는 클러스터 내 로그 관리 대몬(daemon)이 중지되면 해당 대몬(daemon)을 자동으로 다시 시작합니다. 이 개선 사항은 디스크 과다 사용으로 인해 노드가 비정상으로 보이는 위험을 줄입니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용](#)을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

6.9.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.1.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.6.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.23.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.2.0	EMR S3Select 커넥터
emrfs	2.54.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.15.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.15.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-1.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-1.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-1.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-1.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-1.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-1.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.3-amzn-1.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	3.3.3-amzn-1.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-1.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-1.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-1.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.13-amzn-0.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.13-amzn-0.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.13-amzn-0.1	HBase 명령줄 클라이언트
hbase-rest-server	2.4.13-amzn-0.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.13-amzn-0.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.13-amzn-0.1	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-2.1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	3.1.3-amzn-2.1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-2.1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-2.1	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-2.1	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-2.1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-2.1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.12.1-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.12.1-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.12.1-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.12.1-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
iceberg	0.14.1-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.7.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.276-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.276-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.276-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	398-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	398-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	398-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.3.0-amzn-1.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.3.0-amzn-1.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.0-amzn-1.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.0-amzn-1.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.08.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.10.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-0.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.9.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.9.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services SecondaryNamenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.9.1 변경 로그

6.9.1 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-30	릴리스 정보 업데이트	릴리스 정보에 컨트롤 플레인 관련 수정 사항이 여러 개 추가되었습니다.
2023-08-21	문서 게시	Amazon EMR 6.9.1 릴리스 정보 최초 게시됨
2023-08-16	배포 완료	모든 지원되는 리전에 Amazon EMR 6.9.1 전체 배포됨
2023-08-04	초기 릴리스	제한된 상용 리전에 Amazon EMR 6.9.1 최초 배포됨

Amazon EMR 릴리스 6.9.0

6.9.0 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Oozie, Phoenix, Pig, SparkSqoop, PrestoTensorFlow, Tez, 및 애플리케이션이 포함되어 Trino Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
AWS Java용 SDK	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.15
AmazonCloudWatchAgent	-	-	-	-
Delta	2.1.0	-	-	-
Flink	1.15.2	1.15.1	1.15.1	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.13	2.4.12	2.4.12	2.4.4
HCatalog	3.1.3	3.1.3	3.1.3	3.1.3
Hadoop	3.3.3	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.3
Hudi	0.12.1-amzn-0	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.1-amzn-0	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.9.1	1.8.0

	emr-6.9.0	emr-6.8.1	emr-6.8.0	emr-6.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.276	0.273	0.273	0.272
Spark	3.3.0	3.3.0	3.3.0	3.2.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.10.0	2.9.1	2.9.1	2.4.1
Tez	0.10.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	398	388	388	378
Zeppelin	0.10.1	0.10.1	0.10.1	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.10	3.5.7

6.9.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.9.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 릴리스 6.8.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

새로운 기능

- Amazon EMR 릴리스 6.9.0에서는 Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398, Tez 0.10.2를 지원합니다.
- Amazon EMR 릴리스 6.9.0에는 새로운 오픈 소스 애플리케이션, Delta Lake 2.1.0이 포함되어 있습니다.
- Apache Spark용 Amazon Redshift 통합은 Amazon EMR 릴리스 6.9.0 이상에 포함되어 있습니다. 이전의 오픈 소스 도구였던, 이 기본 통합은 Spark 커넥터로, Amazon Redshift와 Amazon Redshift

Serverless에서 데이터를 읽고 쓰는 Apache Spark 애플리케이션을 빌드할 수 있습니다. 자세한 내용은 Amazon EMR에서 Apache Spark용 Amazon Redshift 통합 사용 단원을 참조하십시오.

- Amazon EMR 릴리스 6.9.0은 클러스터 스케일 다운 중에 Amazon S3에 로그를 아카이브하기 위한 지원을 추가합니다. 이전에는 클러스터 종료 중에만 Amazon S3에 로그 파일을 아카이브할 수 있었습니다. 새로운 기능을 통해 클러스터에서 생성된 로그 파일은 노드가 종료된 후에도 Amazon S3에서 계속 유지됩니다. 자세한 내용은 클러스터 로깅 및 디버깅 구성을 참조하세요.
- 이제 Trino에는 장기 실행 쿼리를 지원하기 위해 내결함성 실행 메커니즘이 포함됩니다. 내결함성 실행은 실패한 쿼리 또는 구성 요소 작업을 재시도하여 쿼리 실패를 완화합니다.
- Iceberg, Kinesis 또는 Kafka와 같은 Flink 테이블 소스의 메타데이터 또는 Apache Hive 테이블의 통합된 `BATCH` 및 `STREAM` 처리를 위해 Amazon EMR에서 Apache Flink를 사용할 수 있습니다. AWS Management Console AWS CLI또는 Amazon EMR API를 사용하여 AWS Glue 데이터 카탈로그를 Flink의 메타스토어로 지정할 수 있습니다. 자세한 내용은 Amazon EMR에서 Flink 구성 단원을 참조하십시오.
- 이제 Amazon SageMaker AI Studio를 사용하여 Amazon EMR on EC2 클러스터에서 Apache Spark, Apache Hive 및 Presto 쿼리에 대한 AWS Identity and Access Management (IAM) 런타임 역할 및 AWS Lake Formation기반 액세스 제어를 지정할 수 있습니다. Amazon SageMaker 자세한 내용은 Amazon EMR 단계에 대한 런타임 역할 구성을 참조하세요.

알려진 문제

- Amazon EMR 릴리스 6.9.0의 경우 Trino는 Apache Ranger가 활성화된 클러스터에서 작동하지 않습니다. Ranger와 함께 Trino를 사용해야 하는 경우 지원에 문의하세요.
- Apache Spark용 Amazon Redshift 통합을 사용하고 Parquet 형식의 time, timetz, timestamp 또는 timestamptz(마이크로초 정밀도)를 사용하는 경우 커넥터는 시간 값을 가장 가까운 밀리초 값으로 반올림합니다. 해결 방법으로 텍스트 언로드 형식 `unload_s3_format` 파라미터를 사용합니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가

이 카테고리에 속합니다. `!"#$%&'()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon SageMaker AI Studio에서 Amazon EMR 클러스터에 연결하면 403 금지 응답 코드와 함께 간헐적으로 실패할 수 있습니다. 이 오류는 클러스터에서 IAM 역할을 설정하는 데 60초를 초과하는 경우 발생합니다. 해결 방법으로 Amazon EMR 패치를 설치하여 재시도를 활성화하고 제한 시간을 최소 300초로 늘릴 수 있습니다. 다음 단계를 사용하여 클러스터를 시작할 때 부트스트랩 작업을 적용합니다.

  1. 다음 Amazon S3 URI에서 부트스트랩 스크립트와 RPM 파일을 다운로드합니다.

     ```
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
     s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
     ```

  2. 이전 단계의 파일을 사용자가 소유한 Amazon S3 버킷에 업로드합니다. 버킷은 클러스터를 시작하려는 AWS 리전 동일한에 있어야 합니다.

  3. EMR 클러스터를 시작할 때 다음 부트스트랩 작업을 포함합니다. *`bootstrap_URI`* 및 *`RPM_URI`*를 Amazon S3의 해당되는 URI로 바꿉니다.

     ```
     --bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
     ```

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- Apache Flink는 네이티브 S3 FileSystem 및 Hadoop FileSystem 커넥터를 제공합니다. 이 커넥터를 사용하면 애플리케이션에서 FileSink를 생성하고 Amazon S3에 데이터를 쓸 수 있습니다. 이 FileSink는 다음 두 가지 예외 중 하나로 인해 실패합니다.

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
 supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError:
 org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/RetryPolicy;
                                        at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-yarn-common-3.3.3-amzn-0.jar:?]
```

해결 방법으로 Amazon EMR 패치를 설치할 수 있습니다. 그러면 Flink에서 위 문제를 해결할 수 있습니다. 클러스터를 시작할 때 부트스트랩 작업을 적용하려면 다음 단계를 완료합니다.

1. flink-rpm을 사용자의 Amazon S3 버킷에 다운로드합니다. RPM 경로는 `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`입니다.
2. 다음 URI를 사용하여 Amazon S3에서 부트스트랩 스크립트와 RPM 파일을 다운로드합니다. 를 클러스터를 시작하려는 AWS 리전 *regionName*로 바꿉니다.

```
s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-rpms.sh
```

3. Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. Amazon EMR 6.8.0 및 6.9.0에서 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

   Amazon EMR 6.10.0에서 이 문제를 해결하려면 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications`의 값을 `false`로 설정합니다. Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 기본적으로 구성이 `false`로 설정되어 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 릴리스 6.9.0 이상의 경우 Log4j 라이브러리를 사용하는 Amazon EMR에서 설치한 모든 구성 요소는 Log4j 버전 2.17.1 이상을 사용합니다.
- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 6.9.0에서는 Spark SQL을 사용하여 데이터를 읽을 때 Apache Hudi를 사용한 Lake Formation 기반 액세스 제어에 대한 제한된 지원을 추가했습니다. Spark SQL을 사용하는 SELECT 쿼리가 지원되며 열 수준의 액세스 제어로 제한됩니다. 자세한 내용은 Hudi 및 Lake Formation을 참조하세요.
- Amazon EMR 6.9.0을 사용하여 노드 레이블을 활성화한 상태로 Hadoop 클러스터를 생성하면 YARN 지표 API에서 기본 파티션이 아닌 모든 파티션에서 집계된 정보를 반환합니다. 자세한 내용은 YARN-11414를 참조하세요.
- Amazon EMR 릴리스 6.9.0에서 Trino를 버전 398(Java 17 사용)로 업데이트했습니다. Amazon EMR 6.8.0에서 이전에 지원되던 Trino 버전은 Java 11에서 실행되는 Trino 388이었습니다. 이 변경 사항에 대한 자세한 내용은 Trino 블로그에서 Trino updates to Java 17을 참조하세요.
- 이 릴리스에서는 Apache BigTop 및 Amazon EMR on EC2 클러스터 시작 시퀀스 간 타이밍 시퀀스 불일치 문제를 수정했습니다. 이 타이밍 시퀀스 불일치는 시스템에서 둘 이상의 작업을 적절한 순서로 수행하는 대신 동시에 수행하려고 할 때 발생합니다. 따라서 특정 클러스터 구성에서 인스턴스 시작 제한 시간을 초과하고 클러스터 시작 시간이 느려지는 문제가 발생했습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

> **Note**
>
> 이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, `ListReleaseLabels` API 또는 `list-release-labels` CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20230404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 1월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelease Label(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

6.9.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
delta	2.1.0	Delta Lake는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.6.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.23.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.2.0	EMR S3Select 커넥터
emrfs	2.54.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.15.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.15.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.3.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.3.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.3.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.3.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.3.3-amzn-1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.3.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.3.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	3.3.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.3.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.3.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.3.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.13-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.13-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.13-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.4.13-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.13-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.13-amzn-0	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	3.1.3-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-2	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.12.1-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.12.1-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.12.1-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.12.1-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
iceberg	0.14.1-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.7.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	6.0.0-SNAPSHOT	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	6.0.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.276-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.276-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.276-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	398-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	398-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	398-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.3.0-amzn-1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.3.0-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.0-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.0-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.08.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.10.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.10.2-amzn-0	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.9.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.9.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
delta-defaults	Delta의 delta-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.

분류	설명	재구성 작업
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j2	Livy log4j2.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.

분류	설명	재구성 작업
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)

분류	설명	재구성 작업
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-lakeformation	Presto의 lakeformation.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-delta	Trino의 delta.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-exchange-manager	Trino의 exchange-manager.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.9.0 변경 로그

6.9.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-30	릴리스 정보 업데이트	타이밍 시퀀스 불일치 문제에 대한 수정 사항 추가됨
2023-08-21	릴리스 정보 업데이트	Hadoop 3.3.3에서 알려진 문제가 추가되었습니다.
2023-07-26	업데이트	새 OS 릴리스 레이블 `2.0.20230612.0` 및 `2.0.20230628.0` .
2022-12-13	릴리스 정보 업데이트됨	SageMaker AI를 사용한 런타임에 대한 기능 및 알려진 문제 추가
2022-11-29	릴리스 정보 및 설명서를 업데이트함	Apache Spark용 Amazon Redshift 통합에 대한 기능 추가됨
2022-11-23	릴리스 정보 업데이트됨	Log4j 항목 제거됨
2022-11-18	배포 완료	모든 지원되는 리전에 Amazon EMR 6.9 전체 배포됨
2022-11-18	문서 게시	Amazon EMR 6.9 릴리스 정보 최초 게시됨
2022-11-14	초기 릴리스	제한된 상용 리전에 Amazon EMR 6.9 최초 배포됨

Amazon EMR 릴리스 6.8.1

6.8.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Oozie, Phoenix, Pig, Spark, PrestoSqoop, TensorFlowTez, Trino, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
AWS Java용 SDK	1.12.170	1.12.170	1.12.170	1.12.170
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.15	2.12.15	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.15.1	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.12	2.4.4	2.4.4

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
HCatalog	3.1.3	3.1.3	3.1.3	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.3	3.1.2
Hudi	0.11.1-amzn-0	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	0.14.0-amzn-0	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.9.1	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.273	0.273	0.272	0.267
Spark	3.3.0	3.3.0	3.2.1	3.2.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.9.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.8.1	emr-6.8.0	emr-6.7.0	emr-6.6.0
Trino (PrestoSQL)	388	388	378	367
Zeppelin	0.10.1	0.10.1	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.10	3.5.7	3.5.7

6.8.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.8.1에 대한 정보가 포함됩니다. 변경 사항은 6.8.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.8.1 변경 로그 섹션을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` 값이 `false`로 설정됩니다.

  이 수정 사항은 YARN-9608 때문에 발생한 문제를 해결하지만, Managed Scaling이 활성화된 클러스터에서는 셔플 데이터 손실로 인해 Hive 작업이 실패할 수 있습니다. 이번 릴리스에서는 Hive 워크로드에 대해 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data`도 설정하여 이러한 위험을 완화했습니다. 이 구성은 Amazon EMR 릴리스 6.11.0 이상에서만 사용할 수 있습니다.
- 지표 수집기는 인스턴스 그룹 구성을 사용하는 클러스터의 프라이머리 노드를 장애 조치한 후 컨트롤 플레인에 지표를 전송하지 않습니다.
- 이 릴리스에서는 지표 수집기 엔드포인트에 대한 실패한 HTTP 요청의 재시도가 제거되었습니다.
- 이 릴리스에는 다시 시작한 후 고가용성 클러스터를 장애 상태에서 복구할 수 있도록 지원하는 변경 사항이 포함되어 있습니다.
- 이 릴리스에서는 사용자가 생성한 UID가 클 경우 오버플로 예외가 발생하는 문제를 수정했습니다.
- 이 릴리스에서는 Amazon EMR 재구성 프로세스와 관련된 제한 시간 초과 문제를 수정했습니다.

- 이번 릴리스에서는 재구성 실패로 인해 관련 없는 다른 프로세스가 중단될 수 있는 문제를 방지합니다.
- 이 릴리스에는 보안 수정 사항이 포함되어 있습니다.
- 이 릴리스에서는 Amazon EMR을 사용하여 Spark에서 워크로드를 실행하는 클러스터가 `contains`, `startsWith`, `endsWith`, `like`에서 잘못된 결과를 자동으로 수신할 수 있는 문제를 수정했습니다. 이 문제는 Amazon EMR Hive3 Metastore Server(HMS)에 메타데이터가 있는 파티셔닝된 필드에서 표현식을 사용할 때 발생합니다.
- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 [HIVE-20703](https://issues.apache.org/jira/browse/HIVE-20703) 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 [HIVE-22269](https://issues.apache.org/jira/browse/HIVE-22269) 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.
- HDFS를 스크래치 디렉터리로 사용하고 작은 파일 병합을 활성화하며 테이블에 정적 파티션 경로가 포함된 Hive에서 데이터 손실이 발생할 수 있습니다.
- 이 릴리스에서는 ETL 작업 종료 시 작은 작은 파일 병합(기본적으로 비활성화됨)이 활성화된 경우 Hive의 성능 문제를 수정했습니다.
- 이 릴리스에서는 사용자 정의 함수(UDF)가 없을 때 발생하는 Glue 측 제한 관련 문제를 수정했습니다.
- 이 릴리스에서는 YARN 서비스 해제 시 로그 푸셔가 컨테이너 로그를 S3로 푸시하기 전에 노드 로그 집계 서비스에서 컨테이너 로그를 삭제하는 문제를 수정했습니다.
- 이 릴리스에서는 HBase에 대한 지속적 스토어 파일 추적에서 압축된 파일 및 아카이브된 파일의 처리를 수정합니다.
- 이 릴리스에서는 `spark-defaults.conf`에서 `spark.yarn.heterogeneousExecutors.enabled` 구성의 기본 `true` 값을 설정할 때 Spark 성능에 영향을 주던 문제를 수정했습니다.
- 이 릴리스에서는 Reduce 작업이 셔플 데이터를 읽지 못하는 문제를 수정했습니다. 이 문제로 인해 Hive 쿼리가 실패하고 메모리 손상 오류가 발생했습니다.
- 이 릴리스에서는 노드 대치 중에 HDFS Namenode(NN) 서비스가 안전 모드에서 멈춘 경우 노드 프로비저너에 오류가 발생하는 문제를 수정했습니다.
- 이 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항을 통해 크기 조정 작업의 한 번 실패로 인해 클러스터

크기 조정이 무기한 중단될 위험이 줄어듭니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.

- 이 릴리스에서는 클러스터가 클러스터의 HDFS 복제 계수 설정 미만으로 코어 노드의 스케일 다운을 시도하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이는 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 로컬 인스턴스 스토리지의 열린 파일 핸들 및 관련 프로세스에서 현재 사용 중인 모든 로그를 식별하도록 로그 관리 대몬(daemon)이 업그레이드되었습니다. 이 업그레이드를 통해 Amazon EMR은 로그를 Amazon S3에 아카이브한 후 파일을 올바르게 삭제하고 스토리지 공간을 확보할 수 있습니다.
- 이 릴리스에는 로컬 클러스터 파일 시스템에서 사용되지 않는 빈 단계 디렉터리를 삭제하는 로그 관리 대몬(daemon) 개선 사항이 포함되어 있습니다. 빈 디렉터리가 너무 많으면 Amazon EMR 대몬(daemon)의 성능이 저하되고 디스크가 과도하게 사용될 수 있습니다.
- 이 릴리스에서는 여러 프라이머리 노드를 포함하는 클러스터에서 프라이머리 노드 중 하나를 복제하여 엣지 노드를 생성할 때 발생할 수 있는 문제를 수정했습니다. 복제된 엣지 노드로 인해 스케일 다운 작업이 지연되거나 프라이머리 노드의 메모리 사용률이 높아질 수 있습니다. EMR 클러스터와 통신하기 위해 엣지 노드를 생성하는 방법에 대한 자세한 내용은 GitHub의 `aws-samples` 리포지토리에서 Edge Node Creator를 참조하세요.
- 이 릴리스에서는 Amazon EMR이 재부팅 후 Amazon EBS 볼륨을 인스턴스에 다시 마운트하는 데 사용하는 자동화 프로세스를 개선합니다.
- 이 릴리스에서는 Amazon EMR이 Amazon CloudWatch에 게시하는 Hadoop 지표에 간헐적으로 간격이 발생하는 문제를 수정했습니다.
- 이 릴리스에서는 클러스터의 노드 제외 목록이 포함된 YARN 구성 파일에 대한 업데이트가 디스크 과다 사용으로 인해 중단되는 EMR 클러스터 관련 문제를 수정했습니다. 업데이트가 불완전하면 향후 클러스터 스케일 다운 작업에 방해가 됩니다. 이 릴리스에서는 클러스터를 정상 상태로 유지하고 조정 작업을 예상대로 수행하도록 보장합니다.
- 이 릴리스에서는 MR 클러스터의 추가 로그 폴더를 모니터링하도록 클러스터 내 로그 관리 대몬(daemon)이 개선되었습니다. 이 개선 사항은 디스크 과다 사용 시나리오를 최소화합니다.
- 이 릴리스에서는 클러스터 내 로그 관리 대몬(daemon)이 중지되면 해당 대몬(daemon)을 자동으로 다시 시작합니다. 이 개선 사항은 디스크 과다 사용으로 인해 노드가 비정상으로 보이는 위험을 줄입니다.
- 이번 릴리스에서는 클러스터 스케일 다운 중에 Amazon S3에 로그 아카이브에 대한 지원이 추가되었습니다. 이전에는 클러스터 종료 중에만 Amazon S3에 로그 파일을 아카이브할 수 있었습니다. 새로운 기능을 통해 클러스터에서 생성된 로그 파일은 노드가 종료된 후에도 Amazon S3에서 계속 유지됩니다. 자세한 내용은 클러스터 로깅 및 디버깅 구성을 참조하세요.

- 이 릴리스에서는 부트스트랩 작업에 대한 Amazon S3 URI가 포트 번호로 끝날 때(예: `a.b.c.d:4345`) 발생한 문제를 수정합니다. Amazon EMR이 이러한 URI를 잘못 구문 분석했기 때문에 관련 부트스트랩 작업이 모두 실패했습니다.
- 이 릴리스에서는 Apache BigTop 및 Amazon EMR on EC2 클러스터 시작 시퀀스 간 타이밍 시퀀스 불일치 문제를 수정했습니다. 이 타이밍 시퀀스 불일치는 시스템에서 둘 이상의 작업을 적절한 순서로 수행하는 대신 동시에 수행하려고 할 때 발생합니다. 따라서 특정 클러스터 구성에서 인스턴스 시작 제한 시간을 초과하고 클러스터 시작 시간이 느려지는 문제가 발생했습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 캐나다(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

6.8.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.22.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.53.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.15.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.15.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-8.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-8.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-8.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-8.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-8.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-8.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-8.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-8.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-8.1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	3.2.1-amzn-8.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-8.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.12-amzn-0.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.12-amzn-0.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.12-amzn-0.1	HBase 명령줄 클라이언트
hbase-rest-server	2.4.12-amzn-0.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.12-amzn-0.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.12-amzn-0.1	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-1.1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-1.1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-1.1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-1.1	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-1.1	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	3.1.3-amzn-1.1	하듑 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-1.1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.11.1-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.11.1-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.11.1-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.11.1-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하듑 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	0.14.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.7.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.2	Spark-3용 Apache Phoenix 커넥터
phoenix-query-server	5.1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.273.3-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.273.3-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.273.3-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.

구성 요소	버전	설명
trino-coordinator	388-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	388-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	388-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.3.0-amzn-0.1	Spark 명령줄 클라이언트.
spark-history-server	3.3.0-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.0-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.0-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.06.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tensorflow	2.9.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.8.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.8.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.

분류	설명	재구성 작업
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.

분류	설명	재구성 작업
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.8.1 변경 로그

6.8.1 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-30	릴리스 정보 업데이트	릴리스 정보에 컨트롤 플레인 관련 수정 사항이 여러 개 추가되었습니다.

날짜	이벤트	설명
2023-08-21	문서 게시	Amazon EMR 6.8.1 릴리스 정보 최초 게시됨
2023-08-16	배포 완료	모든 지원되는 리전에 Amazon EMR 6.8.1 전체 배포됨
2023-08-04	초기 릴리스	제한된 상용 리전에 Amazon EMR 6.8.1 최초 배포됨

Amazon EMR 릴리스 6.8.0

6.8.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Oozie, Phoenix, Pig, Spark, PrestoSqoop, TensorFlowTez, Trino, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
AWS Java용 SDK	1.12.170	1.12.170	1.12.170	1.12.31
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
Scala	2.12.15	2.12.15	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.15.1	1.14.2	1.14.2	1.14.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.12	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.3	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.3	3.1.2	3.1.2
Hudi	0.11.1-amzn-0	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	0.14.0-amzn-0	0.13.1-amzn-0	0.13.1	0.12.0
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.9.1	1.8.0	1.8.0	1.8.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2

	emr-6.8.0	emr-6.7.0	emr-6.6.0	emr-6.5.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.273	0.272	0.267	0.261
Spark	3.3.0	3.2.1	3.2.0	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.9.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	388	378	367	360
Zeppelin	0.10.1	0.10.0	0.10.0	0.10.0
ZooKeeper	3.5.10	3.5.7	3.5.7	3.5.7

6.8.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.8.0에 대한 정보가 포함됩니다. 변경 사항은 6.7.0에 관련됩니다.

새로운 기능

- Amazon EMR 단계 기능은 이제 Apache Livy 엔드포인트와 JDBC 및 ODBC 클라이언트를 지원합니다. 자세한 내용은 Amazon EMR 단계에 대한 런타임 역할 구성을 참조하세요.
- Amazon EMR 릴리스 6.8.0은 Apache HBase 릴리스 2.4.12와 함께 제공됩니다. 이 HBase 릴리스에서는 HBase 테이블을 아카이브 및 삭제할 수 있습니다. Amazon S3 아카이브 프로세스는 모든 테이블 파일의 이름을 아카이브 디렉터리로 바꿉니다. 이 프로세스는 비용이 많이 들고 시간이 오래 걸릴 수 있습니다. 이제 아카이브 프로세스를 건너뛰고 대용량 테이블을 빠르게 삭제할 수 있습니다. 자세한 내용은 HBase 쉘 사용 단원을 참조하십시오.

알려진 문제

- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. Amazon EMR 6.8.0 및 6.9.0에서 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 6.10.0에서 이 문제를 해결하려면 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications`의 값을 `false`로 설정합니다. Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 기본적으로 구성이 `false`로 설정되어 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 쉘을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 릴리스 6.8.0은 Apache Hudi 0.11.1과 함께 제공되지만, Amazon EMR 6.8.0 클러스터는 Hudi 0.12.0의 오픈 소스 `hudi-spark3.3-bundle_2.12`와도 호환됩니다.
- Amazon EMR 릴리스 6.8.0은 Apache Spark 3.3.0과 함께 제공됩니다. 이번 Spark 릴리스에서는 Apache Log4j 2 및 `log4j2.properties` 파일을 사용하여 Spark 프로세스에서 Log4j를 구성합니다. 클러스터에서 Spark를 사용하거나 사용자 지정 구성 파라미터를 사용하여 EMR 클러스터를 생성하고 Amazon EMR 릴리스 6.8.0으로 업그레이드하려는 경우 Apache Log4j 2의 새로운 `spark-log4j2` 구성 분류 및 키 형식으로 마이그레이션해야 합니다. 자세한 내용은 Apache Log4j 1.x에서 Log4j 2.x로 마이그레이션 단원을 참조하십시오.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

> **Note**
>
> 이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나,

ListReleaseLabels API 또는 list-release-labels CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20230404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230119.1	4.14.301	2023년 2월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2022 210.1	4.14.301	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2022 004.0	4.14.294	2022년 11월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220912.1	4.14.291	2022년 9월 6일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

알려진 문제

- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Spark 3.2.0에서는 기본적으로 `spark.hadoopRDD.ignoreEmptySplits`를 `true`로 설정하기 때문입니다. 해결 방법으로 `spark.hadoopRDD.ignoreEmptySplits`를 명시적으로 `false`로 설정합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.

  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
  ```

릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

6.8.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.2	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.7.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.22.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.53.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.15.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.15.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-8	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-8	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-8	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-8	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-8	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-8	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-8	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	3.2.1-amzn-8	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-8	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-8	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-8	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.12-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.12-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.12-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.4.12-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.12-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.12-amzn-0	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.3-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	3.1.3-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-1	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.11.1-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.11.1-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.11.1-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.11.1-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
iceberg	0.14.0-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.9.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.7.0	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.2	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	5.1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.273.3-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.273.3-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.273.3-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	388-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	388-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	388-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.3.0-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.3.0-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.3.0-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.3.0-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.06.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.9.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.10	ZooKeeper 명령줄 클라이언트.

6.8.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.8.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j2	Spark의 log4j2.properties 파일에서 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

6.8.0 변경 로그

6.8.0 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-08-21	업데이트	Hadoop 3.3.3에서 알려진 문제가 추가되었습니다.

날짜	이벤트	설명
2023-07-26	업데이트	새 OS 릴리스 레이블 `2.0.20230612.0` 및 `2.0.20230628.0` .
2022-09-06	배포 완료	모든 지원되는 리전에 Amazon EMR 6.8 전체 배포됨
2022-09-06	최초 게시	Amazon EMR 6.8 릴리스 정보 최초 게시됨
2022-08-31	초기 릴리스	제한된 상용 리전에 Amazon EMR 6.8 출시됨

Amazon EMR 릴리스 6.7.0

6.7.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Oozie, Phoenix, Pig, Spark, PrestoSqoop, TensorFlowTez, Trino, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
AWS Java용 SDK	1.12.170	1.12.170	1.12.31	1.12.31
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.15	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.0	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.4.4
HCatalog	3.1.3	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.3	3.1.2	3.1.2	3.1.2
Hudi	0.11.0-amzn-0	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	0.13.1-amzn-0	0.13.1	0.12.0	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0

	emr-6.7.0	emr-6.6.0	emr-6.5.0	emr-6.4.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.1.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.272	0.267	0.261	0.254.1
Spark	3.2.1	3.2.0	3.1.2	3.1.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	378	367	360	359
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.5.7

6.7.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.7.0에 대한 정보가 포함됩니다. 변경 사항은 6.6.0에 관련됩니다.

최초 릴리스 날짜: 2022년 7월 15일

새로운 기능

- Amazon EMR은 이제 Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272, Trino 0.378을 지원합니다.
- Amazon EMR on EC2 클러스터에 대한 EMR 단계(Spark, Hive)를 통해 IAM 역할 및 Lake Formation 기반 액세스 제어를 지원합니다.

- Apache Ranger 지원 클러스터에서 Apache Spark 데이터 정의 명령문을 지원합니다. 이제 Apache Ranger 지원 클러스터에서 Apache Hive 메타데이터를 읽고 쓰는 Trino 애플리케이션에 대한 지원이 포함됩니다. 자세한 내용은 Enable federated governance using Trino and Apache Ranger on Amazon EMR을 참조하세요.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 이스라엘(텔아비브), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20230404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230119.1	4.14.301	2023년 2월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2022 004.0	4.14.294	2022년 11월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220912.1	4.14.291	2022년 10월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20220719.0	4.14.287	2022년 8월 10일	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220606.1	4.14.281	2022년 7월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

알려진 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 셸을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 잘못된 `Hbase.compat.version`을 사용하므로 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Spark 3.2.0에서는 기본적으로 `spark.hadoopRDD.ignoreEmptySplits`를 `true`로 설정하기 때문입니다. 해결 방법으로 `spark.hadoopRDD.ignoreEmptySplits`를 명시적으로 `false`로 설정합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.
- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
  ```
- Amazon EMR 6.7 이하에서 실행되는 클러스터에서는 `GetClusterSessionCredentials` API가 지원되지 않습니다.

6.7.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.6.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.22.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.52.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.

구성 요소	버전	설명
flink-client	1.14.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.14.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-7	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-7	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-7	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-7	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-7	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-7	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.2.1-amzn-7	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-7	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-7	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-7	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-7	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.4-amzn-3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.4-amzn-3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.4-amzn-3	HBase 명령줄 클라이언트
hbase-rest-server	2.4.4-amzn-3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.4-amzn-3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.4-amzn-3	Apache HBase 클러스터용 복구 도구.

구성 요소	버전	설명
hcatalog-client	3.1.3-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.3-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	3.1.3-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.3-amzn-0	Hive 명령줄 클라이언트
hive-hbase	3.1.3-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	3.1.3-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.3-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.11.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.11.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.11.0-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.11.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	0.13.1-amzn-0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.0.194	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.2	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	5.1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.272-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.272-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.272-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	378-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	378-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	378-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.2.1-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.2.1-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.2.1-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.2.1-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.02.0-amzn-1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.7	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.7	ZooKeeper 명령줄 클라이언트.

6.7.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.7.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.6.0

6.6.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hudi, Hive, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Oozie, Phoenix, Pig, Spark, PrestoSqoop, TensorFlowTez, Trino, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
AWS Java용 SDK	1.12.170	1.12.31	1.12.31	1.11.977
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.0	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.4.4	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.10.1-amzn-0	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0

	emr-6.6.0	emr-6.5.0	emr-6.4.0	emr-6.3.1
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	0.13.1	0.12.0	-	-
JupyterEn terpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.1.2	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.261	0.254.1	0.245.1
Spark	3.2.0	3.1.2	3.1.2	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQ L)	367	360	359	350
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.5.7	3.4.14

6.6.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.6.0에 대한 정보가 포함됩니다. 변경 사항은 6.5.0에 관련됩니다.

최초 릴리스 날짜: 2022년 5월 9일

설명서 업데이트 날짜: 2022년 6월 15일

새로운 기능

- Amazon EMR 6.6에서는 Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367, PrestoDB 0.267을 지원합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 캐나다(중부), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.2023 404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230119.1	4.14.301	2023년 2월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221004.0	4.14.294	2022년 11월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220912.1	4.14.291	2022년 10월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20220805.0	4.14.287	2022년 8월 30일	`us-west-1`

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220719.0	4.14.287	2022년 8월 10일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220426.0	4.14.281	2022년 6월 10일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220406.1	4.14.275	2022년 5월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

- Amazon EMR 릴리스 6.6 이상에서 Log4j 1.x 및 Log4j 2.x를 사용하는 애플리케이션은 각각 Log4j 1.2.17 이상 및 Log4j 2.17.1 이상을 사용하도록 업그레이드되었으며, CVE 문제를 완화하기 위해 부트스트랩 작업을 사용할 필요가 없습니다.
- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.
- Amazon EMR 5.32.0 및 6.5.0부터 Apache Spark의 동적 실행기 크기 조정이 기본적으로 활성화됩니다. 이 기능을 켜거나 끄기 위해 `spark.yarn.heterogeneousExecutors.enabled` 구성 파라미터를 사용할 수 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR에서는 EMR 기본 AMI 옵션을 사용하고 Apache Hadoop, Apache Spark, Apache Hive와 같은 일반적인 애플리케이션만 설치하는 클러스터에 대해 클러스터 시작 시간을 평균 최대 80초까지 줄일 수 있습니다.

알려진 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 셸을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 잘못된 `Hbase.compat.version`을 사용하므로 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Spark 3.2.0에서는 기본적으로 `spark.hadoopRDD.ignoreEmptySplits`를 `true`로 설정하기 때문입니다. 해결 방법으로 `spark.hadoopRDD.ignoreEmptySplits`를 명시적으로 `false`로 설정합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.
- Trino 장기 실행 클러스터에서 Amazon EMR 6.6.0을 사용하면 Trino jvm.config의 가비지 수집 로깅 파라미터를 통해 가비지 수집 로그에서 더 나은 인사이트를 얻을 수 있습니다. 이 변경으로 많은 가비지 수집 로그가 launcher.log(/var/log/trino/launcher.log) 파일에 추가됩니다. Amazon EMR 6.6.0에서 Trino 클러스터를 실행하는 경우 클러스터를 며칠 동안 실행한 후에 추가된 로그로 인해 노드의 디스크 공간이 부족해질 수 있습니다.

  이 문제의 해결 방법은 Amazon EMR 6.6.0에서 클러스터를 생성하거나 복제하는 동안 아래 스크립트를 부트스트랩 작업으로 실행하여 jvm.config에서 가비지 수집 로깅 파라미터를 비활성화하는 것입니다.

  ```
  #!/bin/bash
    set -ex
    PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
    sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
  ```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.

  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-
actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-
actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

6.6.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.5.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.20.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.50.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.14.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.14.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-6	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	3.2.1-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.4-amzn-2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.4-amzn-2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.4-amzn-2	HBase 명령줄 클라이언트
hbase-rest-server	2.4.4-amzn-2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.4-amzn-2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hbase-operator-tools	2.4.4-amzn-2	Apache HBase 클러스터용 복구 도구.
hcatalog-client	3.1.2-amzn-7	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	3.1.2-amzn-7	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-7	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-7	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-7	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-7	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-7	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.10.1-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.10.1-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.10.1-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.10.1-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
iceberg	0.13.1	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	11.0.194	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-connectors	5.1.2	Spark-3용 Apache Phoenix 커넥터

구성 요소	버전	설명
phoenix-query-server	5.1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.267-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.267-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.267-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	367-amzn-0	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	367-amzn-0	여러 쿼리를 실행하는 서비스.
trino-client	367-amzn-0	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.2.0-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.2.0-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.2.0-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.2.0-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	22.02.0-amzn-0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.7	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.7	ZooKeeper 명령줄 클라이언트.

6.6.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.6.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts PhoenixQu eryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)

분류	설명	재구성 작업
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-iceberg	Trino의 iceberg.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.5.0

6.5.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hudi, Hive, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Oozie, Phoenix, Pig, Spark, PrestoSqoop, TensorFlowTez, Trino, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
AWS Java용 SDK	1.12.31	1.12.31	1.11.977	1.11.977
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.0	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.4.4	2.2.6	2.2.6
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.9.0-amzn-1	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0

	emr-6.5.0	emr-6.4.0	emr-6.3.1	emr-6.3.0
Hue	4.9.0	4.9.0	4.9.0	4.9.0
Iceberg	0.12.0	-	-	-
JupyterEn terpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	5.1.2	5.1.2	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.261	0.254.1	0.245.1	0.245.1
Spark	3.1.2	3.1.2	3.1.1	3.1.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQ L)	360	359	350	350
Zeppelin	0.10.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.5.7	3.4.14	3.4.14

6.5.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.5.0에 대한 정보가 포함됩니다. 변경 사항은 6.4.0에 관련됩니다.

최초 릴리스 날짜: 2022년 1월 20일

릴리스 날짜: 2022년 3월 21일

새로운 기능

- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.
- Amazon EMR 5.32.0 및 6.5.0부터 Apache Spark의 동적 실행기 크기 조정이 기본적으로 활성화됩니다. 이 기능을 켜거나 끄기 위해 `spark.yarn.heterogeneousExecutors.enabled` 구성 파라미터를 사용할 수 있습니다.
- 대규모 분석 데이터 세트를 위한 Apache Iceberg 오픈 테이블 형식에 대한 지원.
- ranger-trino-plugin 2.0.1-amzn-1 지원
- toree 0.5.0에 대한 지원

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 6.5 릴리스 버전은 이제 Apache Iceberg 0.12.0을 지원하며, Apache Spark용 EMR 런타임, Presto용 Amazon EMR 런타임, Apache Hive용 Amazon EMR 런타임을 통해 런타임을 개선했습니다.
- Apache Iceberg는 Amazon S3에서 대형 데이터 세트를 위한 오픈 테이블 형식이며, 대형 테이블, 원자성 커밋, 동시 쓰기, SQL 호환 테이블 진화 등에서 빠른 쿼리 성능을 제공합니다. EMR 6.5에서 Iceberg 테이블 형식으로 Apache Spark 3.1.2를 사용할 수 있습니다.
- Apache Hudi 0.9에서는 Spark SQL DDL 및 DML 지원을 추가했습니다. 이를 통해 SQL 문만 사용하여 Hudi 테이블을 생성하고 업서트할 수 있습니다. Apache Hudi 0.9에는 쿼리 측 및 작성기 측 성능 개선도 포함되어 있습니다.
- Apache Hive용 Amazon EMR 런타임은 준비 작업 중에 이름 바꾸기 작업을 제거하여 Amazon S3의 Apache Hive 성능을 향상시키고, 테이블 복구에 사용되는 메타스토어 검사(MSCK) 명령의 성능을 개선합니다.

알려진 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 셸을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 잘못된 `Hbase.compat.version`을 사용하므로 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- 고가용성(HA)의 Hbase 번들 클러스터에서 기본 볼륨 크기 및 인스턴스 유형으로 프로비저닝하지 못합니다. 이 문제의 해결 방법은 루트 볼륨 크기를 늘리는 것입니다.
- Apache Oozie에서 Spark 작업을 사용하려면 Oozie `workflow.xml` 파일에 다음 구성을 추가해야 합니다. 그렇지 않으면 Oozie가 시작하는 Spark 실행기의 클래스 경로에서 Hadoop 및 EMRFS와 같은 여러 중요한 라이브러리가 누락됩니다.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

6.5.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.4.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.19.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.48.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.14.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.14.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-5	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	3.2.1-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.4-amzn-1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.4-amzn-1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.4-amzn-1	HBase 명령줄 클라이언트
hbase-rest-server	2.4.4-amzn-1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.4-amzn-1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-6	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-6	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	3.1.2-amzn-6	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-6	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-6	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-6	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-6	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.9.0-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.9.0-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.9.0-amzn-1	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.9.0-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.9.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
iceberg	0.12.0	Apache Iceberg는 방대한 분석 데이터 세트를 위한 오픈 테이블 형식입니다.

구성 요소	버전	설명
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	5.1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.261-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.261-amzn-0	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
presto-client	0.261-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	360	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	360	여러 쿼리를 실행하는 서비스.
trino-client	360	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.1.2-amzn-1	Spark 명령줄 클라이언트.
spark-history-server	3.1.2-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.1.2-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.1.2-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	0.4.1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.

구성 요소	버전	설명
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.7	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.7	ZooKeeper 명령줄 클라이언트.

6.5.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.5.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services

분류	설명	재구성 작업
		ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
iceberg-defaults	Iceberg의 iceberg-defaults.conf 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.4.0

6.4.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Oozie, Phoenix, , Pig, Presto, Sqoop, SparkTensorFlow, Tez, Trino및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
AWS Java용 SDK	1.12.31	1.11.977	1.11.977	1.11.880
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.4.4	2.2.6	2.2.6	2.2.6-amzn-0

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.8.0-amzn-0	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	5.1.2	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.254.1	0.245.1	0.245.1	0.238.3
Spark	3.1.2	3.1.1	3.1.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.4.0	emr-6.3.1	emr-6.3.0	emr-6.2.1
Trino (PrestoSQL)	359	350	350	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.5.7	3.4.14	3.4.14	3.4.14

6.4.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.4.0에 대한 정보가 포함됩니다. 변경 사항은 6.3.0에 관련됩니다.

최초 릴리스 날짜: 2021년 9월 20일

릴리스 날짜: 2022년 3월 21일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.12.31
- CloudWatch Sink 버전 2.2.0
- DynamoDB 커넥터 버전 4.16.0
- EMRFS 버전 2.47.0
- Amazon EMR Goodies 버전 3.2.0
- Amazon EMR Kinesis 커넥터 버전 3.5.0
- Amazon EMR Record Server 버전 2.1.0
- Amazon EMR 스크립트 버전 2.5.0
- Flink 버전 1.13.1
- Ganglia 버전 3.7.2
- AWS Glue Hive 메타스토어 클라이언트 버전 3.3.0
- Hadoop 버전 3.2.1-amzn-4
- HBase 버전 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog 버전 3.1.2-amzn-5

- Hive 버전 3.1.2-amzn-5
- Hudi 버전 0.8.0-amzn-0
- Hue 버전 4.9.0
- Java JDK 버전 Corretto-8.302.08.1(빌드 1.8.0_302-b08)
- JupyterHub 버전 1.4.1
- Livy 버전 0.7.1-incubating
- MXNet 버전 1.8.0
- Oozie 버전 5.2.1
- Phoenix 버전 5.1.2
- Pig 버전 0.17.0
- Presto 버전 0.254.1-amzn-0
- Trino 버전 359
- Apache Ranger KMS(멀티 마스터 투명 암호화) 버전 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.2.0
- SageMaker Spark SDK 버전 1.4.1
- Scala 버전 2.12.10(OpenJDK 64비트 서버 VM, Java 1.8.0_282)
- Spark 버전 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop 버전 1.4.7
- TensorFlow 버전 2.4.1
- tez 버전 0.9.2
- Zeppelin 버전 0.9.0
- Zookeeper 버전 3.5.7
- 커넥터 및 드라이버: DynamoDB 커넥터 4.16.0

새로운 특성

- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하

기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.

- Apache Ranger 지원 Amazon EMR 클러스터에서는 Apache Spark SQL을 사용하여 `INSERT INTO`, `INSERT OVERWRITE`, `ALTER TABLE`을 통해 Apache Hive 메타스토어 테이블에 데이터를 삽입하거나 업데이트할 수 있습니다. Spark SQL에서 ALTER TABLE을 사용하는 경우 파티션 위치는 테이블 위치의 하위 디렉터리여야 합니다. Amazon EMR은 현재 파티션 위치가 테이블 위치와 다른 파티션으로의 데이터 삽입을 지원하지 않습니다.
- PrestoSQL은 Trino로 이름이 바뀌었습니다.
- Hive: LIMIT 절에 언급된 레코드 수를 가져오는 즉시 쿼리 실행을 중지함으로써 LIMIT 절을 사용하는 간단한 SELECT 쿼리의 실행이 가속화됩니다. 간단한 SELECT 쿼리는 GROUP BY 및 ORDER BY 절이 없는 쿼리 또는 reducer 스테이지가 없는 쿼리입니다. 예를 들어 `SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`입니다.

Hudi 동시성 제어

- Hudi는 이제 낙관적 동시성 제어(OCC)를 지원합니다. 이 기능을 UPSERT 및 INSERT와 같은 쓰기 작업과 함께 활용하면 여러 작성기에서 동일한 Hudi 테이블로 변경할 수 있습니다. 이는 파일 수준 OCC이므로 변경 사항이 충돌하지 않는 한 두 커밋(또는 작성기)이 동일한 테이블에 쓸 수 있습니다. 자세한 내용은 Hudi concurrency control을 참조하세요.
- Amazon EMR 클러스터에는 Zookeeper가 설치되어 있고, 이를 OCC의 잠금 공급자로 활용할 수 있습니다. 이 기능을 더 쉽게 사용할 수 있도록 Amazon EMR 클러스터에는 다음 속성이 미리 구성되어 있습니다.

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

OCC를 활성화하려면 Hudi 작업 옵에서 다음 속성을 구성하거나 Amazon EMR 구성 API를 사용하여 클러스터 수준에서 다음 속성을 구성해야 합니다.

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Hudi 모니터링: Hudi 지표를 보고하기 위한 Amazon CloudWatch 통합

- Amazon EMR은 Amazon CloudWatch에 Hudi 지표를 게시하는 기능을 지원합니다. 다음과 같은 필수 구성을 설정하여 이 기능을 활성화합니다.

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- 변경할 수 있는 선택적 Hudi 구성은 다음과 같습니다.

설정	설명	값
hoodie.metrics.cloudwatch.report.period.seconds	Amazon CloudWatch에 지표를 보고하는 빈도(초)	기본값은 60초이며, Amazon CloudWatch에서 제공하는 기본 1분 분해능에 적합합니다.
hoodie.metrics.cloudwatch.metric.prefix	각 지표 이름에 추가할 접두사	기본값은 비어 있습니다(접두사 없음).
hoodie.metrics.cloudwatch.namespace	지표가 게시되는 Amazon CloudWatch 네임스페이스	기본값은 Hudi입니다.
hoodie.metrics.cloudwatch.maxDatumsPerRequest	Amazon CloudWatch에 대한 1개의 요청에 포함할 수 있는 최대 데이텀 수	기본값은 20이며, Amazon CloudWatch 기본값과 동일합니다.

Amazon EMR Hudi 구성 지원 및 개선 사항

- 이제 고객은 EMR 구성 API 및 재구성 기능을 활용하여 클러스터 수준에서 Hudi 구성을 설정할 수 있습니다. Spark, Hive 등과 같은 다른 애플리케이션과 마찬가지로 /etc/hudi/conf/hudi-defaults.conf를 통해 새로운 파일 기반 구성 지원이 도입되었습니다. EMR은 사용자 환경을 개선하기 위해 몇 가지 기본값을 구성합니다.

  \- `hoodie.datasource.hive_sync.jdbcurl` 은 클러스터 Hive 서버 URL로 구성되며, 더 이상 지정하지 않아도 됩니다. 이전에 Amazon EMR 마스터 IP를 지정해야 했던 Spark 클러스터 모드에서 작업을 실행할 때 특히 유용합니다.

  \- HBase 특정 구성으로, Hudi에서 HBase 인덱스를 사용하는 데 유용합니다.

- Zookeeper 잠금 제공자 특정 구성으로, 동시성 제어에서 설명한 대로, 이를 통해 낙관적 동시성 제어(OCC)를 더 쉽게 사용할 수 있습니다.

• 통과해야 하는 구성 수를 줄이고 가능한 경우 자동으로 추론할 수 있도록 추가 변경 사항이 도입되었습니다.

- `partitionBy` 키워드를 사용하여 파티션 열을 지정할 수 있습니다.

- Hive Sync를 활성화하면 더 이상 `HIVE_TABLE_OPT_KEY`, `HIVE_PARTITION_FIELDS_OPT_KEY`, `HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`를 전달하지 않아도 됩니다. 이러한 값은 Hudi 테이블 이름 및 파티션 필드에서 추론할 수 있습니다.

- `KEYGENERATOR_CLASS_OPT_KEY`는 반드시 전달하지 않아도 되며, `SimpleKeyGenerator` 및 `ComplexKeyGenerator`의 보다 간단한 사례에서 추론할 수 있습니다.

Hudi Caveats

• Hudi는 Merge on Read(MoR) 및 부트스트랩 테이블에 대해 Hive에서 벡터화된 실행을 지원하지 않습니다. 예를 들어 `hive.vectorized.execution.enabled`가 true로 설정된 경우 Hudi 실시간 테이블의 `count(*)`에서 오류가 발생합니다. 해결 방법으로 `hive.vectorized.execution.enabled`를 `false`로 설정하여 벡터화된 읽기를 비활성화할 수 있습니다.

• 다중 작성기 지원은 Hudi 부트스트랩 기능과 호환되지 않습니다.

• Flink Streamer 및 Flink SQL은 이 릴리스의 실험 기능입니다. 이러한 기능은 프로덕션 배포에 사용하지 않는 것이 좋습니다.

변경 사항, 향상된 기능 및 해결된 문제

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

• 이전에는 다중 마스터 클러스터에서 리소스 관리자를 수동으로 다시 시작하면 Zookeeper와 같은 Amazon EMR 클러스터 내 대몬(daemon)이 Zookeeper znode 파일에서 이전에 해제되거나 손실된 노드를 모두 다시 로드했습니다. 이로 인해 특정 상황에서 기본 한도가 초과되었습니다. 이제 Amazon EMR은 Zookeeper 파일에서 1시간 넘게 경과된 해제되거나 손실된 노드 레코드를 제거하고 내부 한도를 늘렸습니다.

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- Apache YARN 타임라인 서버 버전 1 및 1.5 성능 문제를 해결하도록 클러스터 구성

  Apache YARN 타임라인 서버 버전 1 및 1.5는 매우 자주 사용되는 대규모 EMR 클러스터, 특히 Amazon EMR의 기본 설정인 `yarn.resourcemanager.system-metrics-publisher.enabled=true`에서 성능 문제를 일으킬 수 있습니다. 오픈 소스 YARN 타임라인 서버 v2는 YARN 타임라인 서버 확장성과 관련된 성능 문제를 해결했습니다.

  이 문제의 다른 해결 방법은 다음과 같습니다.

  - yarn-site.xml 파일에서 yarn.resourcemanager.system-metrics-publisher.enabled=false를 구성합니다.
  - 아래 설명에 따라 클러스터를 생성할 때 이 문제를 해결하는 수정 사항을 활성화합니다.

  다음 Amazon EMR 릴리스에는 이 YARN 타임라인 서버 성능 문제에 대한 수정 사항이 포함되어 있습니다.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  위에 지정된 Amazon EMR 릴리스에서 수정 사항을 활성화하려면 `aws emr create-cluster` 명령 파라미터(`--configurations file://./configurations.json`)를 사용하여 전달되는 구

성 JSON 파일에서 이러한 속성을 `true`로 설정합니다. 또는 재구성 콘솔 UI를 사용하여 수정 사항을 활성화할 수도 있습니다.

configurations.json 파일 내용 예제:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- WebHDFS 및 HttpFS 서버는 기본적으로 비활성화됩니다. Hadoop 구성(`dfs.webhdfs.enabled`)을 사용하여 WebHDFS를 다시 활성화할 수 있습니다. HttpFS 서버는 `sudo systemctl start hadoop-httpfs`를 사용하여 시작할 수 있습니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.
- Hive: 마지막 작업에 HDFS의 스크래치 디렉터리를 사용할 수 있게 함으로써 쓰기 쿼리 성능이 개선되었습니다. 최종 작업의 임시 데이터는 Amazon S3 대신 HDFS에 기록되며, 데이터는 Amazon S3 디바이스 사이가 아닌 HDFS에서 최종 테이블 위치(Amazon S3)로 이동되므로 성능이 향상됩니다.
- Hive: Glue 메타스토어 파티션 정리를 통해 쿼리 컴파일 시간이 최대 2.5배 개선되었습니다.
- 기본적으로 내장된 UDF를 Hive에서 Hive Metastore Server로 전달하면 Glue는 제한된 표현식 연산자만 지원하므로 내장된 UDF의 하위 세트만 Glue 메타스토어로 전달됩니다. `hive.glue.partition.pruning.client=true`를 설정하면 모든 파티션 정리가 클라이언트 측에서 수행됩니다. `hive.glue.partition.pruning.server=true`를 설정하면 모든 파티션 정리가 서버 측에서 수행됩니다.

알려진 문제

- Apache Hadoop HttpFS 서버는 기본적으로 비활성화되어 있기 때문에 Amazon EMR 6.4.0에서는 Hue 쿼리가 작동하지 않습니다. Amazon EMR 6.4.0에서 Hue를 사용하려면 `sudo systemctl start hadoop-httpfs`를 사용하여 Amazon EMR 프라이머리 노드에서 HttpFS 서버를 수동으로 시작하거나 Amazon EMR 단계를 사용합니다.
- Livy 사용자 위장과 함께 사용되는 Amazon EMR Notebooks 기능은 기본적으로 HttpFS 가 비활성화되었기 때문에 작동하지 않습니다. 이 경우 EMR 노트북은 Livy 가장 기능이 활성화된 클러스터에 연결할 수 없습니다. 해결 방법은 클러스터에 EMR 노트북을 연결하기 전에 `sudo systemctl start hadoop-httpfs`를 사용하여 HttpFS 서버를 시작하는 것입니다.
- Amazon EMR 버전 6.4.0에서 Phoenix는 Phoenix 커넥터 구성 요소를 지원하지 않습니다.
- Apache Oozie에서 Spark 작업을 사용하려면 Oozie `workflow.xml` 파일에 다음 구성을 추가해야 합니다. 그렇지 않으면 Oozie가 시작하는 Spark 실행기의 클래스 경로에서 Hadoop 및 EMRFS와 같은 여러 중요한 라이브러리가 누락됩니다.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

6.4.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을

위해 설치됩니다. 이는 일반적으로 emr 또는 aws로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 myapp-component라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 2.2-amzn-2로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.3.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.18.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.47.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.13.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
flink-jobmanager-config	1.13.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.2.1-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.4.4-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.4.4-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.4.4-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.4.4-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.4.4-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-5	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	3.1.2-amzn-5	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-5	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-5	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-5	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-5	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-5	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.8.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.8.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-trino	0.8.0-amzn-0	Hudi와 함께 Trnio를 실행하기 위한 번들 라이브러리.
hudi-spark	0.8.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.9.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	5.1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.254.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.254.1-amzn-0	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
presto-client	0.254.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
trino-coordinator	359	trino-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
trino-worker	359	여러 쿼리를 실행하는 서비스.
trino-client	359	Trino 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Trino 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.1.2-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.1.2-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.1.2-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.1.2-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	0.4.1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.

구성 요소	버전	설명
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.5.7	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.5.7	ZooKeeper 명령줄 클라이언트.

6.4.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.4.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services

분류	설명	재구성 작업
		ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Resta rts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hb ase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
trino-log	Trino의 log.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-config	Trino의 config.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-password-authenticator	Trino의 password-authenticator.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-env	Trino의 trino-env.sh 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-node	Trino의 node.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-blackhole	Trino의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-cassandra	Trino의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-hive	Trino의 hive.properties 파일에서 값을 변경합니다.	Restarts Trino-Server (for Trino)
trino-connector-jmx	Trino의 jmx.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-kafka	Trino의 kafka.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-localfile	Trino의 localfile.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
trino-connector-memory	Trino의 memory.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mongodb	Trino의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-mysql	Trino의 mysql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-postgresql	Trino의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-raptor	Trino의 raptor.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redis	Trino의 redis.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-redshift	Trino의 redshift.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpch	Trino의 tpch.properties 파일에서 값을 변경합니다.	Not available.
trino-connector-tpcds	Trino의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.3.1

6.3.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHub, JupyterEnterpriseGatewayLivy, MXNet, Oozie, Phoenix, Pig, PrestoSQL, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
AWS Java용 SDK	1.11.977	1.11.977	1.11.880	1.11.880
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.245.1	0.238.3	0.238.3
Spark	3.1.1	3.1.1	3.0.1	3.0.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.3.1	emr-6.3.0	emr-6.2.1	emr-6.2.0
Trino (PrestoSQL)	350	350	343	343
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.3.1 릴리스 정보

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사

항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

6.3.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.2.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.18.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.46.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.12.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.12.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-3.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-3.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-3.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-3.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-3.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-3.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-3.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-3.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-3.1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	3.2.1-amzn-3.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-3.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.6-amzn-1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.6-amzn-1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.2.6-amzn-1	HBase 명령줄 클라이언트
hbase-rest-server	2.2.6-amzn-1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.6-amzn-1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-4	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-4	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	3.1.2-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.7.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.7.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-prestosql	0.7.0-amzn-0	Hudi와 함께 PrestoSQL을 실행하기 위한 번들 라이브러리.
hudi-spark	0.7.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.9.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.2.2	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.

구성 요소	버전	설명
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.245.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.245.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.245.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
prestosql-coordinator	350	prestosql-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
prestosql-worker	350	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
prestosql-client	350	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.1.1-amzn-0.1	Spark 명령줄 클라이언트.
spark-history-server	3.1.1-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.1.1-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.1.1-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	0.4.1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.9.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.3.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.3.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
prestosql-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	PrestoSQL의 node.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-blackhole	PrestoSQL의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-cassandra	PrestoSQL의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-hive	PrestoSQL의 hive.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL의 jmx.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-kafka	PrestoSQL의 kafka.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-localfile	PrestoSQL의 localfile.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
prestosql-connector-memory	PrestoSQL의 memory.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mongodb	PrestoSQL의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mysql	PrestoSQL의 mysql.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-postgresql	PrestoSQL의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-raptor	PrestoSQL의 raptor.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redis	PrestoSQL의 redis.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redshift	PrestoSQL의 redshift.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpch	PrestoSQL의 tpch.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpcds	PrestoSQL의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.3.0

6.3.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Oozie, Phoenix, , Pig, Presto, Spark, PrestoSQLTez, SqoopTensorFlow, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
AWS Java용 SDK	1.11.977	1.11.880	1.11.880	1.11.828
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.7.0-amzn-0	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.238.3	0.238.3	0.232
Spark	3.1.1	3.0.1	3.0.1	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.3.0	emr-6.2.1	emr-6.2.0	emr-6.1.1
Trino (PrestoSQL)	350	343	343	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.3.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.3.0에 대한 정보가 포함됩니다. 변경 사항은 6.2.0에 관련됩니다.

최초 릴리스 날짜: 2021년 5월 12일

최종 업데이트 날짜: 2021년 8월 9일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.11.977
- CloudWatch Sink 버전 2.1.0
- DynamoDB 커넥터 버전 4.16.0
- EMRFS 버전 2.46.0
- Amazon EMR Goodies 버전 3.2.0
- Amazon EMR Kinesis 커넥터 버전 3.5.0
- Amazon EMR Record Server 버전 2.0.0
- Amazon EMR 스크립트 버전 2.5.0
- Flink 버전 1.12.1
- Ganglia 버전 3.7.2
- AWS Glue Hive 메타스토어 클라이언트 버전 3.2.0
- Hadoop 버전 3.2.1-amzn-3
- HBase 버전 2.2.6-amzn-1
- HBase-operator-tools 1.0.0
- HCatalog 버전 3.1.2-amzn-0

- Hive 버전 3.1.2-amzn-4
- Hudi 버전 0.7.0-amzn-0
- Hue 버전 4.9.0
- Java JDK 버전 Corretto-8.282.08.1(빌드 1.8.0_282-b08)
- JupyterHub 버전 1.2.0
- Livy 버전 0.7.0-incubating
- MXNet 버전 1.7.0
- Oozie 버전 5.2.1
- Phoenix 버전 5.0.0
- Pig 버전 0.17.0
- Presto 버전 0.245.1-amzn-0
- PrestoSQL 버전 350
- Apache Ranger KMS(멀티 마스터 투명 암호화) 버전 2.0.0
- ranger-plugins 2.0.1-amzn-0
- ranger-s3-plugin 1.1.0
- SageMaker Spark SDK 버전 1.4.1
- Scala 버전 2.12.10(OpenJDK 64비트 서버 VM, Java 1.8.0_282)
- Spark version 3.1.1-amzn-0
- spark-rapids 0.4.1
- Sqoop 버전 1.4.7
- TensorFlow 버전 2.4.1
- tez 버전 0.9.2
- Zeppelin 버전 0.9.0
- Zookeeper 버전 3.4.14
- 커넥터 및 드라이버: DynamoDB 커넥터 4.16.0

새로운 특성

- Amazon EMR은 공유 데이터 레이크에 대한 액세스를 쉽게 관리할 수 있는 Amazon S3의 기능인 Amazon S3 액세스 포인트를 지원합니다. Amazon S3 액세스 포인트 별칭을 사용하면 Amazon EMR에서 대규모 데이터 액세스를 간소화할 수 있습니다. Amazon EMR을 사용할 수 있는 모든 리전

에서 추가 비용 없이 모든 버전의 Amazon EMR에서 Amazon S3 액세스 포인트를 사용할 수 있습니다. AWS Amazon S3 액세스 포인트 및 액세스 포인트 별칭에 대한 자세한 내용은 Amazon S3 사용 설명서에서 액세스 포인트에 버킷 스타일 별칭 사용을 참조하세요.

- 새 `DescribeReleaseLabel` 파라미터 및 `ListReleaseLabel` API 파라미터에서는 Amazon EMR 릴리스 레이블 세부 정보를 제공합니다. API 요청이 실행되는 리전에서 사용 가능한 릴리스를 프로그래밍 방식으로 나열하고 특정 Amazon EMR 릴리스 레이블에서 사용할 수 있는 애플리케이션을 나열할 수 있습니다. 릴리스 레이블 파라미터는 Spark와 같은 특정 애플리케이션을 지원하는 Amazon EMR 릴리스도 나열합니다. 이 정보는 Amazon EMR 클러스터를 프로그래밍 방식으로 시작하는 데 사용할 수 있습니다. 예를 들어 `ListReleaseLabel` 결과에서 최신 릴리스 버전을 사용하여 클러스터를 시작할 수 있습니다. 자세한 내용은 Amazon EMR API 참조에서 DescribeReleaseLabel 및 ListReleaseLabels를 참조하세요.
- Amazon EMR 6.3.0에서는 Apache Ranger와 기본적으로 통합되는 클러스터를 시작할 수 있습니다. Apache Ranger는 Hadoop 플랫폼 전반에서 포괄적인 데이터 보안을 지원, 모니터링 및 관리할 수 있는 오픈 소스 프레임워크입니다. 자세한 내용은 Apache Ranger를 참조하세요. 기본 통합을 통해 자체 Apache Ranger를 이용해 Amazon EMR에서 세분화된 데이터 액세스 제어를 적용할 수 있습니다. Amazon EMR 관리 안내서에서 Amazon EMR과 Apache Ranger 통합을 참조하세요.
- 범위가 지정된 관리형 정책: AWS 모범 사례에 맞게 Amazon EMR은 더 이상 사용되지 않을 정책을 대체하기 위해 v2 EMR 범위 기본 관리형 정책을 도입했습니다. Amazon EMR 관리형 정책을 참조하세요.
- 인스턴스 메타데이터 서비스(IMDS) V2 지원 상태: Amazon EMR 6.2 이상의 Amazon EMR 구성 요소는 모든 IMDS 직접 호출에 IMDSv2를 사용합니다. 애플리케이션 코드에서 IMDS를 직접 호출하는 경우, IMDSv1 및 IMDSv2를 모두 사용하거나, 추가 보안을 위해 IMDSv2만 사용하도록 IMDS를 구성할 수 있습니다. 이전 Amazon EMR 6.x 릴리스에서 IMDSv1을 비활성화하면 클러스터를 시작하지 못합니다.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.

- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Spark 3.1에서는 Spark SQL UI 설명 모드 기본값이 `extended`에서 `formatted`로 변경되었습니다. Amazon EMR은 이를 다시 `extended`로 되돌려 Spark SQL UI에 논리적 계획 정보를 포함합니다. `spark.sql.ui.explainMode`를 `formatted`로 설정하여 이를 되돌릴 수 있습니다.
- 다음 커밋은 Spark 마스터 브랜치에서 백포트되었습니다.

  - [SPARK-34752][빌드] Jetty를 9.4.37로 변경하여 CVE-2020-27223을 해결합니다.

  - [SPARK-34534] FetchShuffleBlocks를 사용하여 블록을 가져올 때 blockIds 순서를 수정합니다.

  - [SPARK-34681] [SQL] 왼쪽을 같지 않음 조건으로 빌드할 때 전체 외부 셔플 해시 조인에 대한 버그를 수정했습니다.

  - [SPARK-34497] [SQL] 기본 제공 JDBC 연결 제공업체를 수정하여 JVM 보안 컨텍스트 변경 사항을 복원합니다.
- Nvidia Spark Rapids 플러그인과의 상호 운용성을 개선하기 위해 적응형 쿼리 실행이 비활성화된 상태에서 Nvidia Spark RAPID를 사용할 때 동적 파티션 정리가 트리거되지 않는 문제를 해결하는 해결 방법을 추가했습니다. RAPIDS issue #1378 및 RAPIDS issue ##1386을 참조하세요. 새 구성 `spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`에 대한 자세한 내용은 내용은 RAPIDS issue ##1386을 참조하세요.
- 오픈 소스 Spark 3.1의 파일 출력 커미터 기본 알고리즘이 v2 알고리즘에서 v1 알고리즘으로 변경되었습니다. 자세한 내용은 이 Amazon EMR optimizing Spark performance - dynamic partition pruning을 참조하세요.

- Amazon EMR은 성능 저하를 방지하기 위해 이전 Amazon EMR 6.x 릴리스에서 사용된 기본값인 v2 알고리즘으로 되돌렸습니다. 오픈 소스 Spark 3.1 동작을 복원하려면 `spark.hadoop.mapreduce.fileoutputcommitter.algorithm.version`을 1로 설정합니다. 파일 출력 커밋 알고리즘 v2의 작업 커밋이 원자성 작업이 아니므로 오픈 소스 Spark에서 이와 같이 변경됩니다. 이로 인해 일부 경우에 출력 데이터 수정 문제가 발생할 수 있습니다. 하지만 알고리즘 v1의 작업 커밋도 원자성 작업이 아닙니다. 일부 시나리오에서는 이름을 바꾸기 전에 작업 커밋에 수행된 삭제가 포함됩니다. 이로 인해 자동 데이터 수정 문제가 발생할 수 있습니다.
- 이전 Amazon EMR 릴리스에서 Managed Scaling 문제를 수정하고 애플리케이션 장애 비율을 크게 줄일 수 있도록 개선했습니다.
- 각 새 클러스터에 AWS Java SDK 번들을 설치했습니다. 개별 구성 요소 jar 대신 모든 서비스 SDK와 해당 종속 항목을 포함하는 단일 jar입니다. 자세한 내용은 Java SDK Bundled Dependency를 참조하세요.

알려진 문제

- Amazon EMR 6.3.0 및 6.2.0 프라이빗 서브넷 클러스터의 경우 Ganglia 웹 UI에 액세스할 수 없습니다. '액세스 거부됨(403)' 오류가 발생합니다. Spark, Hue, JupyterHub, Zeppelin, Livy, Tez와 같은 다른 웹 UI는 정상적으로 작동합니다. 퍼블릭 서브넷 클러스터에서 Ganglia 웹 UI 액세스도 정상적으로 작동합니다. 이 문제를 해결하려면 `sudo systemctl restart httpd`를 사용하여 프라이머리 노드에서 httpd 서비스를 다시 시작합니다. 이 문제는 Amazon EMR 6.4.0에서 수정되었습니다.
- AWS Glue 데이터 카탈로그가 활성화되면 Spark를 사용하여 null 문자열 위치 URI가 있는 AWS Glue DB에 액세스하지 못할 수 있습니다. 이 문제는 이전 Amazon EMR 릴리스에서 발생하지만, SPARK-31709(https://issues.apache.org/jira/browse/SPARK-31709)에서는 더 많은 사례에 이를 적용합니다. 예를 들어 위치 URI가 null 문자열인 기본 AWS Glue DB 내에 테이블을 생성할 때는 "빈 문자열에서 경로를 생성할 수 없습니다."라는 메시지와 함께 `spark.sql("CREATE TABLE mytest (key string) location '/table_path';")` 실패합니다. 이 문제를 해결하려면 AWS Glue 데이터베이스의 위치 URI를 수동으로 설정한 다음 Spark를 사용하여 이러한 데이터베이스 내에 테이블을 생성합니다.
- Amazon EMR 6.3.0에서 PrestoSQL은 버전 343에서 버전 350으로 업그레이드되었습니다. 이 버전 변경과 관련된 오픈 소스의 두 가지 보안 관련 변경 사항이 있습니다. 테이블, 스키마 또는 세션 속성 규칙이 정의되지 않은 경우로 파일 기반 카탈로그 액세스 제어가 `deny`에서 `allow`로 변경됩니다. 또한 파일 기반 시스템 액세스 제어는 카탈로그 규칙이 정의되지 않은 파일을 지원하도록 변경되었습니다. 이 경우 카탈로그에 대한 모든 액세스가 허용됩니다.

  자세한 내용은 Release 344 (9 Oct 2020)를 참조하세요.

- Hadoop 사용자 디렉터리(/home/hadoop)는 모든 사용자가 읽을 수 있습니다. Hive 같은 프레임워크에서 읽기 액세스를 허용하는 Unix 755(drwxr-xr-x) 디렉터리 권한을 보유합니다. /home/hadoop 및 하위 디렉터리에 파일을 넣을 수 있지만 중요한 정보를 보호하려면 해당 디렉터리의 권한을 알고 있어야 합니다.
- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

     `LimitNOFILE=65536`

     `LimitNPROC=65536`
  2. InstanceController 다시 시작

     `$ sudo systemctl daemon-reload`

     `$ sudo systemctl restart instance-controller`

  부트스트랩 작업(BA)을 사용하여 ulimit 설정

  클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

  ```
  #!/bin/bash
  ```

```
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon Linux 또는 Amazon Linux 2 Amazon Machine Image(AMI)를 실행하는 EMR 클러스터는 기본 Amazon Linux 동작을 사용하며, 재부팅이 필요한 중요한 핵심 커널 업데이트를 자동으로 다운로드하여 설치하지는 않습니다. 이는 기본 Amazon Linux AMI를 실행하는 다른 Amazon EC2 인스턴스와 동일한 동작입니다. Amazon EMR 릴리스가 출시된 후 재부팅이 필요한 새로운 Amazon Linux 소프트웨어 업데이트(예: 커널, NVIDIA, CUDA 업데이트)를 사용할 수 있게 되면 기본 AMI를 실행하는 EMR 클러스터 인스턴스는 해당 업데이트를 자동으로 다운로드하여 설치하지 않습니다. 커널 업데이트를 받으려면 최신 Amazon Linux AMI를 사용하도록 Amazon EMR AMI를 사용자 지정할 수 있습니다.

- Apache Oozie에서 Spark 작업을 사용하려면 Oozie `workflow.xml` 파일에 다음 구성을 추가해야 합니다. 그렇지 않으면 Oozie가 시작하는 Spark 실행기의 클래스 경로에서 Hadoop 및 EMRFS와 같은 여러 중요한 라이브러리가 누락됩니다.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.

- 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

6.3.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-notebook-env	1.2.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.18.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.1.0	EMR S3Select 커넥터
emrfs	2.46.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.12.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.12.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	3.2.1-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-3	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.6-amzn-1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.6-amzn-1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	2.2.6-amzn-1	HBase 명령줄 클라이언트
hbase-rest-server	2.2.6-amzn-1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.6-amzn-1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-4	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-4	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.7.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.7.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hudi-prestosql	0.7.0-amzn-0	Hudi와 함께 PrestoSQL을 실행하기 위한 번들 라이브러리.
hudi-spark	0.7.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.9.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.2.2	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.245.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.245.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.245.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
prestosql-coordinator	350	prestosql-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
prestosql-worker	350	여러 쿼리를 실행하는 서비스.
prestosql-client	350	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.1.1-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.1.1-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.1.1-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.1.1-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	0.4.1	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.3.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.3.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS services Namenode, Datanode, and ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.

분류	설명	재구성 작업
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
prestosql-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)

분류	설명	재구성 작업
prestosql-node	PrestoSQL의 node.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-blackhole	PrestoSQL의 blackhole .properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-cassandra	PrestoSQL의 cassandra .properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-hive	PrestoSQL의 hive.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL의 jmx.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-kafka	PrestoSQL의 kafka.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-localfile	PrestoSQL의 localfile.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-memory	PrestoSQL의 memory.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mongodb	PrestoSQL의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mysql	PrestoSQL의 mysql.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-postgresql	PrestoSQL의 postgresql.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
prestosql-connector-raptor	PrestoSQL의 raptor.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redis	PrestoSQL의 redis.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redshift	PrestoSQL의 redshift.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpch	PrestoSQL의 tpch.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpcds	PrestoSQL의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.

분류	설명	재구성 작업
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie and HiveServer2.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.

분류	설명	재구성 작업
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.2.1

6.2.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, PrestoSQL, Presto, Spark, Sqoop, TensorFlow, Tez및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전

- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
AWS Java용 SDK	1.11.880	1.11.880	1.11.828	1.11.828
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.6-amzn-0	2.2.5	2.2.5
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.6.0-amzn-1	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0

	emr-6.2.1	emr-6.2.0	emr-6.1.1	emr-6.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.238.3	0.238.3	0.232	0.232
Spark	3.0.1	3.0.1	3.0.0	3.0.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	343	338	338
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.2.1 릴리스 정보

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했

습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.

- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.

- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

6.2.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-notebook-env	1.0.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.16.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.0.0	EMR S3Select 커넥터
emrfs	2.44.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.11.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.11.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-2.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-2.1	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	3.2.1-amzn-2.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-2.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-2.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-2.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-2.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-2.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-2.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-2.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-2.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.6-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.6-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	2.2.6-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.2.6-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.6-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-3	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-3	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-3	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-3	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-3	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-3	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-3	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.6.0-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.6.0-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hudi-prestosql	0.6.0-amzn-1	Hudi와 함께 PrestoSQL을 실행하기 위한 번들 라이브러리.
hudi-spark	0.6.0-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.8.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.238.3-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.238.3-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.238.3-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
prestosql-coordinator	343	prestosql-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
prestosql-worker	343	여러 쿼리를 실행하는 서비스.
prestosql-client	343	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.0.1-amzn-0.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.0.1-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.0.1-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.0.1-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	0.2.0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.3.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0-preview1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.2.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.2.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Not available.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Not available.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Not available.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services SecondaryNamenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS ZKFC.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.

분류	설명	재구성 작업
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.

분류	설명	재구성 작업
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
prestosql-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)

분류	설명	재구성 작업
prestosql-node	PrestoSQL의 node.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-blackhole	PrestoSQL의 blackhole .properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-cassandra	PrestoSQL의 cassandra .properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-hive	PrestoSQL의 hive.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL의 jmx.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-kafka	PrestoSQL의 kafka.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-localfile	PrestoSQL의 localfile.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-memory	PrestoSQL의 memory.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mongodb	PrestoSQL의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mysql	PrestoSQL의 mysql.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-postgresql	PrestoSQL의 postgresql.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
prestosql-connector-raptor	PrestoSQL의 raptor.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redis	PrestoSQL의 redis.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redshift	PrestoSQL의 redshift.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpch	PrestoSQL의 tpch.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpcds	PrestoSQL의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.

분류	설명	재구성 작업
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.

분류	설명	재구성 작업
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.2.0

6.2.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, JupyterEnterpriseGateway, HueJupyterHub, Livy, MXNet, Oozie, Phoenix, Pig, PrestoSQL, Presto, Spark, Sqoop, TensorFlow, Tez및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
AWS Java용 SDK	1.11.880	1.11.828	1.11.828	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.12.10
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.6-amzn-0	2.2.5	2.2.5	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.6.0-amzn-1	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1
Hue	4.8.0	4.7.1	4.7.1	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0

	emr-6.2.0	emr-6.1.1	emr-6.1.0	emr-6.0.1
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	0.17.0	-
Presto	0.238.3	0.232	0.232	0.230
Spark	3.0.1	3.0.0	3.0.0	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	-
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	343	338	338	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.2.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.2.0에 대한 정보가 포함됩니다. 변경 사항은 6.1.0에 관련됩니다.

최초 릴리스 날짜: 2020년 12월 9일

최종 업데이트 날짜: 2021년 10월 4일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.11.828

- emr-record-server 버전 1.7.0
- Flink 버전 1.11.2
- Ganglia 버전 3.7.2
- Hadoop 버전 3.2.1-amzn-1
- HBase 버전 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog 버전 3.1.2-amzn-0
- Hive 버전 3.1.2-amzn-3
- Hudi 버전 0.6.0-amzn-1
- Hue 버전 4.8.0
- JupyterHub 버전 1.1.0
- Livy 버전 0.7.0
- MXNet 버전 1.7.0
- Oozie 버전 5.2.0
- Phoenix 버전 5.0.0
- Pig 버전 0.17.0
- Presto 버전 0.238.3-amzn-1
- PrestoSQL 버전 343
- Spark 버전 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow 버전 2.3.1
- Zeppelin 버전 0.9.0-preview1
- Zookeeper 버전 3.4.14
- 커넥터 및 드라이버: DynamoDB 커넥터 4.16.0

새로운 특성

- HBase: 커밋 단계에서 이름 바꾸기를 제거하고 지속적 HFile 추적을 추가했습니다. Amazon EMR 릴리스 안내서에서 지속적 HFile 추적을 참조하세요.
- HBase: Create a config that forces to cache blocks on compaction을 백포트했습니다.

- PrestoDB: 동적 파티션 정리 기능이 개선되었습니다. 규칙 기반 조인 재정렬은 파티셔닝되지 않은 데이터에서 작동합니다.
- 범위가 지정된 관리형 정책: AWS 모범 사례에 맞게 Amazon EMR은 더 이상 사용되지 않을 정책을 대체하기 위해 v2 EMR 범위 기본 관리형 정책을 도입했습니다. Amazon EMR 관리형 정책을 참조하세요.
- 인스턴스 메타데이터 서비스(IMDS) V2 지원 상태: Amazon EMR 6.2 이상의 Amazon EMR 구성 요소는 모든 IMDS 직접 호출에 IMDSv2를 사용합니다. 애플리케이션 코드에서 IMDS를 직접 호출하는 경우, IMDSv1 및 IMDSv2를 모두 사용하거나, 추가 보안을 위해 IMDSv2만 사용하도록 IMDS를 구성할 수 있습니다. 이전 Amazon EMR 6.x 릴리스에서 IMDSv1을 비활성화하면 클러스터를 시작하지 못합니다.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.

- Spark: Spark 런타임의 성능이 개선되었습니다.

알려진 문제

- Amazon EMR 6.2에서는 EMR 6.2.0의 /etc/cron.d/libinstance-controller-java 파일에 잘못된 권한이 설정되어 있습니다. 파일의 권한이 644(-rw-r--r--)여야 하는 경우 645(-rw-r--r-x)입니다. 따라서 Amazon EMR 버전 6.2는 인스턴스 상태 로그를 로깅하지 않으며 /emr/instance-logs 디렉터리는 비어 있습니다. 이 문제는 Amazon EMR 6.3.0 이상에서 수정되었습니다.

  이 문제를 해결하려면 클러스터 시작 시 다음 스크립트를 부트스트랩 작업으로 실행합니다.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- Amazon EMR 6.2.0 및 6.3.0 프라이빗 서브넷 클러스터의 경우 Ganglia 웹 UI에 액세스할 수 없습니다. '액세스 거부됨(403)' 오류가 발생합니다. Spark, Hue, JupyterHub, Zeppelin, Livy, Tez와 같은 다른 웹 UI는 정상적으로 작동합니다. 퍼블릭 서브넷 클러스터에서 Ganglia 웹 UI 액세스도 정상적으로 작동합니다. 이 문제를 해결하려면 `sudo systemctl restart httpd`를 사용하여 프라이머리 노드에서 httpd 서비스를 다시 시작합니다. 이 문제는 Amazon EMR 6.4.0에서 수정되었습니다.
- Amazon EMR 6.2.0에는 httpd에서 계속 실패하여 Ganglia를 사용할 수 없게 되는 문제가 있습니다. 'cannot connect to the server' 오류가 발생합니다. 이 문제를 포함하는 이미 실행 중인 클러스터를 수정하려면 클러스터 프라이머리 노드에 SSH로 연결하고 `/etc/httpd/conf/httpd.conf`에 있는 `httpd.conf` 파일에 `Listen 80` 줄을 추가합니다. 이 문제는 Amazon EMR 6.3.0에서 수정되었습니다.
- 보안 구성을 사용하는 경우 EMR 6.2.0 클러스터에서 HTTPD에 오류가 발생합니다. 이로 인해 Ganglia 웹 애플리케이션 사용자 인터페이스를 사용할 수 없습니다. Ganglia 웹 애플리케이션 사용자 인터페이스에 액세스하려면 클러스터의 프라이머리 노드에 있는 `/etc/httpd/conf/httpd.conf` 파일에 `Listen 80`을 추가합니다. 클러스터 연결에 대한 자세한 내용은 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.

  또한 보안 구성을 사용할 때 EMR Notebooks는 EMR 6.2.0 클러스터와의 연결을 설정하지 못합니다. 노트북은 커널을 나열하지 못하고 Spark 작업을 제출하지 못합니다. 대신 다른 Amazon EMR 버전에서 EMR Notebooks를 사용하는 것이 좋습니다.
- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최

대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`

2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
```

```
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon EMR 6.1.0 및 6.2.0에는 모든 Hudi 삽입, 업서트 및 삭제 작업에 큰 영향을 미칠 수 있는 성능 문제가 포함되어 있습니다. Amazon EMR 6.1.0 또는 6.2.0에서 Hudi를 사용하려는 경우 AWS 지원팀에 문의하여 패치된 Hudi RPM을 받아야 합니다.

- **Important**

  Amazon Linux 또는 Amazon Linux 2 Amazon Machine Image(AMI)를 실행하는 EMR 클러스터는 기본 Amazon Linux 동작을 사용하며, 재부팅이 필요한 중요한 핵심 커널 업데이트를 자동으로 다운로드하여 설치하지는 않습니다. 이는 기본 Amazon Linux AMI를 실행하는 다른 Amazon EC2 인스턴스와 동일한 동작입니다. Amazon EMR 릴리스가 출시된 후 재부팅이 필요한 새로운 Amazon Linux 소프트웨어 업데이트(예: 커널, NVIDIA, CUDA 업데이트)를 사용할 수 있게 되면 기본 AMI를 실행하는 EMR 클러스터 인스턴스는 해당 업데이트를 자동으로 다운로드하여 설치하지 않습니다. 커널 업데이트를 받으려면 최신 Amazon Linux AMI를 사용하도록 Amazon EMR AMI를 사용자 지정할 수 있습니다.

- Amazon EMR 6.2.0 Maven 아티팩트는 게시되지 않습니다. 향후 Amazon EMR 릴리스와 함께 게시됩니다.
- HBase 스토어파일 시스템 테이블을 사용한 지속적 HFile 추적은 HBase 리전 복제 기능을 지원하지 않습니다. HBase 리전 복제에 대한 자세한 내용은 타임라인이 일치하는 고가용성 읽기를 참조하세요.
- Amazon EMR 6.x 및 EMR 5.x Hive 버킷팅 버전 차이

  EMR 5.x는 OOS Apache Hive 2를 사용하지만, EMR 6.x는 OOS Apache Hive 3을 사용합니다. 오픈 소스 Hive2는 버킷팅 버전 1을 사용하는 반면, 오픈 소스 Hive3은 버킷팅 버전 2를 사용합니다. Hive 2(EMR 5.x) 및 Hive 3(EMR 6.x)의 버킷팅 버전 차이는 Hive 버킷팅 해싱 기능이 다르다는 것을 의미합니다. 아래 예제를 참조하십시오.

다음 표는 각각 EMR 6.x 및 EMR 5.x에서 생성된 예제입니다.

```
-- Using following LOCATION in EMR 6.x
CREATE TABLE test_bucketing (id INT, desc STRING)
PARTITIONED BY (day STRING)
CLUSTERED BY(id) INTO 128 BUCKETS
LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

-- Using following LOCATION in EMR 5.x
LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
```

EMR 6.x 및 EMR 5.x 모두에 동일한 데이터를 삽입합니다.

```
INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
```

S3 위치를 확인하는 경우 EMR 6.x(Hive 3) 및 EMR 5.x(Hive 2) 간에 해싱 함수가 다르기 때문에 버킷팅 파일 이름이 다르게 표시됩니다.

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

또한 EMR 6.x에서 Hive CLI를 통해 다음 명령을 실행하여 버전 차이를 확인할 수도 있습니다. 이 경우 버킷팅 버전 2가 반환됩니다.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일

다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

6.2.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.0.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.16.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.0.0	EMR S3Select 커넥터
emrfs	2.44.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.

구성 요소	버전	설명
flink-client	1.11.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.11.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-2	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	3.2.1-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.6-amzn-0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.6-amzn-0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.2.6-amzn-0	HBase 명령줄 클라이언트
hbase-rest-server	2.2.6-amzn-0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.6-amzn-0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-3	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	3.1.2-amzn-3	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-3	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-3	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-3	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-3	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-3	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.6.0-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.6.0-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-prestosql	0.6.0-amzn-1	Hudi와 함께 PrestoSQL을 실행하기 위한 번들 라이브러리.
hudi-spark	0.6.0-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.8.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.238.3-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.238.3-amzn-1	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
presto-client	0.238.3-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
prestosql-coordinator	343	prestosql-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
prestosql-worker	343	여러 쿼리를 실행하는 서비스.
prestosql-client	343	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.0.1-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	3.0.1-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.0.1-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.0.1-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
spark-rapids	0.2.0	GPU로 Apache Spark를 가속화하는 Nvidia Spark RAPIDS 플러그인.

구성 요소	버전	설명
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.3.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0-preview1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.2.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-6.2.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services

분류	설명	재구성 작업
		ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Not available.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Not available.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Not available.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Secondary Namenode, Datanode, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	This classification should not be reconfigured.
hdfs-env	HDFS 환경의 값을 변경합니다.	Restarts Hadoop HDFS ZKFC.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat server.

분류	설명	재구성 작업
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat server.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.	Sets configurations to launch Hive LLAP service.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2, HiveMetastore, and Hive HCatalog-Server. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.

분류	설명	재구성 작업
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoDB)
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
prestosql-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-node	PrestoSQL의 node.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-blackhole	PrestoSQL의 blackhole.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-cassandra	PrestoSQL의 cassandra.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-hive	PrestoSQL의 hive.properties 파일에서 값을 변경합니다.	Restarts Presto-Server (for PrestoSQL)
prestosql-connector-jmx	PrestoSQL의 jmx.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-kafka	PrestoSQL의 kafka.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-localfile	PrestoSQL의 localfile.properties 파일에서 값을 변경합니다.	Not available.

분류	설명	재구성 작업
prestosql-connector-memory	PrestoSQL의 memory.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mongodb	PrestoSQL의 mongodb.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-mysql	PrestoSQL의 mysql.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-postgresql	PrestoSQL의 postgresql.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-raptor	PrestoSQL의 raptor.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redis	PrestoSQL의 redis.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-redshift	PrestoSQL의 redshift.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpch	PrestoSQL의 tpch.properties 파일에서 값을 변경합니다.	Not available.
prestosql-connector-tpcds	PrestoSQL의 tpcds.properties 파일에서 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restart Oozie.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 6.1.1

6.1.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HudiHue, Livy, JupyterHubMXNet, Oozie, Phoenix, Pig, Presto, Spark, PrestoSQL, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS Java용 SDK	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.1.1 릴리스 정보

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.

- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. [공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다](#).

6.1.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.14.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emr-s3-select	2.0.0	EMR S3Select 커넥터
emrfs	2.42.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.11.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-1.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-1.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-1.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-1.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-1.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	3.2.1-amzn-1.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-1.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-1.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-1.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-1.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-1.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.5	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.5	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.2.5	HBase 명령줄 클라이언트
hbase-rest-server	2.2.5	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.5	HBase에 Thrift 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hcatalog-client	3.1.2-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-2	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.2-incubating-amzn-2	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.2-incubating-amzn-2	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hudi-prestosql	0.5.2-incubating-amzn-2	Hudi와 함께 PrestoSQL을 실행하기 위한 번들 라이브러리.
hudi-spark	0.5.2-incubating-amzn-2	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.7.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.6.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.3.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.232	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.232	여러 쿼리를 실행하는 서비스.
presto-client	0.232	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
prestosql-coordinator	338	prestosql-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
prestosql-worker	338	여러 쿼리를 실행하는 서비스.
prestosql-client	338	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.0.0-amzn-0.1	Spark 명령줄 클라이언트.
spark-history-server	3.0.0-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.0.0-amzn-0.1	YARN용 인 메모리 실행 엔진.

구성 요소	버전	설명
spark-yarn-slave	3.0.0-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.1.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0-preview1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.1.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-6.1.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-env	HDFS 환경의 값을 변경합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.

분류	설명
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.

분류	설명
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
prestosql-log	Presto log.properties 파일의 값을 변경합니다.
prestosql-config	Presto config.properties 파일의 값을 변경합니다.
prestosql-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
prestosql-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
prestosql-node	PrestoSQL의 node.properties 파일에서 값을 변경합니다.
prestosql-connector-blackhole	PrestoSQL의 blackhole.properties 파일에서 값을 변경합니다.

분류	설명
prestosql-connector-cassandra	PrestoSQL의 cassandra.properties 파일에서 값을 변경합니다.
prestosql-connector-hive	PrestoSQL의 hive.properties 파일에서 값을 변경합니다.
prestosql-connector-jmx	PrestoSQL의 jmx.properties 파일에서 값을 변경합니다.
prestosql-connector-kafka	PrestoSQL의 kafka.properties 파일에서 값을 변경합니다.
prestosql-connector-localfile	PrestoSQL의 localfile.properties 파일에서 값을 변경합니다.
prestosql-connector-memory	PrestoSQL의 memory.properties 파일에서 값을 변경합니다.
prestosql-connector-mongodb	PrestoSQL의 mongodb.properties 파일에서 값을 변경합니다.
prestosql-connector-mysql	PrestoSQL의 mysql.properties 파일에서 값을 변경합니다.
prestosql-connector-postgresql	PrestoSQL의 postgresql.properties 파일에서 값을 변경합니다.
prestosql-connector-raptor	PrestoSQL의 raptor.properties 파일에서 값을 변경합니다.
prestosql-connector-redis	PrestoSQL의 redis.properties 파일에서 값을 변경합니다.
prestosql-connector-redshift	PrestoSQL의 redshift.properties 파일에서 값을 변경합니다.
prestosql-connector-tpch	PrestoSQL의 tpch.properties 파일에서 값을 변경합니다.

분류	설명
prestosql-connector-tpcds	PrestoSQL의 tpcds.properties 파일에서 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.

분류	설명
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 6.1.0

6.1.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HudiHue, Livy, JupyterHubMXNet, Oozie, Phoenix, Pig, Presto, Spark, PrestoSQL, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS Java용 SDK	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.1.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.1.0에 대한 정보가 포함됩니다. 변경 사항은 6.0.0에 관련됩니다.

최초 릴리스 날짜: 2020년 9월 4일

최종 업데이트 날짜: 2020년 10월 15일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.11.828

- Flink 버전 1.11.0
- Ganglia 버전 3.7.2
- Hadoop 버전 3.2.1-amzn-1
- HBase 버전 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog 버전 3.1.2-amzn-0
- Hive 버전 3.1.2-amzn-1
- Hudi 버전 0.5.2-incubating
- Hue 버전 4.7.1
- JupyterHub 버전 1.1.0
- Livy 버전 0.7.0
- MXNet 버전 1.6.0
- Oozie 버전 5.2.0
- Phoenix 버전 5.0.0
- Presto 버전 0.232
- PrestoSQL 버전 338
- Spark 버전 3.0.0-amzn-0
- TensorFlow 버전 2.1.0
- Zeppelin 버전 0.9.0-preview1
- Zookeeper 버전 3.4.14
- 커넥터 및 드라이버: DynamoDB Connector 4.14.0

새로운 특성

- ARM 인스턴스 유형은 Amazon EMR 버전 5.30.0 및 Amazon EMR 버전 6.1.0부터 지원됩니다.
- M6g 범용 인스턴스 유형은 Amazon EMR 버전 6.1.0 및 5.30.0부터 지원됩니다. 자세한 내용은 Amazon EMR 관리 안내서에서 지원되는 인스턴스 유형을 참조하세요.
- EC2 배치 그룹 기능은 여러 프라이머리 노드 클러스터에 대한 옵션으로 Amazon EMR 버전 5.23.0부터 지원됩니다. 현재 배치 그룹 기능에서는 프라이머리 노드 유형만 지원되며 해당 프라이머리 노드에는 `SPREAD` 전략이 적용됩니다. `SPREAD` 전략은 하드웨어 장애 발생 시 여러 프라이머리 노드의 손실을 방지하기 위해 별도의 기본 하드웨어에 작은 그룹의 인스턴스를 배치합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR과 EC2 배치 그룹 통합을 참조하세요.

- Managed Scaling - Amazon EMR 버전 6.1.0에서는 Amazon EMR Managed Scaling을 활성화하여 워크로드에 따라 클러스터에서 인스턴스 또는 유닛 수를 자동으로 늘리거나 줄일 수 있습니다. Amazon EMR은 클러스터 지표를 지속적으로 평가하여 비용과 속도 측면에서 클러스터를 최적화하는 조정 결정을 내립니다. Amazon EMR 버전 5.30.0 이상(6.0.0 제외)에서도 Managed Scaling을 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 리소스 조정을 참조하세요.
- PrestoSQL 버전 338은 EMR 6.1.0에서 지원됩니다. 자세한 내용은 Presto를 참조하세요.
  - PrestoSQL은 EMR 6.1.0 이상 버전에서만 지원되며, EMR 6.0.0 또는 EMR 5.x에서 지원되지 않습니다.
  - 애플리케이션 이름(`Presto`)은 클러스터에 PrestoDB를 설치하는 데 계속 사용됩니다. 클러스터에 PrestoSQL을 설치하려면 애플리케이션 이름, `PrestoSQL`을 사용합니다.
  - PrestoDB 또는 PrestoSQL 중 하나를 설치할 수 있지만 단일 클러스터에 둘 다 설치할 수는 없습니다. 클러스터를 생성하려고 할 때 PrestoDB와 PrestoSQL을 모두 지정하면 검증 오류가 발생하고 클러스터 생성 요청이 실패합니다.
  - PrestoSQL은 단일 마스터 클러스터 및 멀티 마스터 클러스터 모두에서 지원됩니다. 멀티 마스터 클러스터에서 PrestoSQL 또는 PrestoDB를 실행하려면 외부 Hive 메타스토어가 필요합니다. 여러 프라이머리 노드를 포함하는 EMR 클러스터에서 지원되는 애플리케이션을 참조하세요.
- Docker를 통해 Apache Hadoop 및 Apache Spark에서의 ECR 자동 인증 지원: Spark 사용자는 Docker Hub와 Amazon Elastic Container Registry(Amazon ECR)의 도커 이미지를 사용하여 환경 및 라이브러리 종속 항목을 정의할 수 있습니다.

  Docker를 구성하고 Amazon EMR 6.x를 사용하여 Docker와 함께 Spark 애플리케이션을 실행합니다.
- EMR에서 Apache Hive ACID 트랜잭션 지원: Amazon EMR 6.1.0에서는 Hive ACID 트랜잭션에 대한 지원을 추가하였고, 데이터베이스의 ACID 속성을 준수합니다. 이 기능을 사용하면 Amazon Simple Storage Service(S3)의 데이터를 사용하여 Hive 관리형 테이블에서 `INSERT`, `UPDATE`, `DELETE`, 및 `MERGE` 작업을 실행할 수 있습니다. 이는 스트리밍 수집, 데이터 복원, MERGE를 사용한 대량 업데이트, 느리게 변화하는 크기 등과 같은 사용 사례와 관련된 주요 기능입니다. 구성 예제 및 사용 사례를 비롯한 자세한 내용은 Amazon EMR supports Apache Hive ACID transactions를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Apache Flink는 EMR 6.0.0에서 지원되지 않지만 EMR 6.1.0에서 Flink 1.11.0으로 지원됩니다. 이 버전은 Hadoop 3을 공식적으로 지원하는 Flink의 첫 번째 버전입니다. Apache Flink 1.11.0 Release Announcement를 참조하세요.
- Ganglia는 기본 EMR 6.1.0 패키지 번들에서 제거되었습니다.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한

보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`
2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
```

```
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- > **Important**
  >
  > Amazon EMR 6.1.0 및 6.2.0에는 모든 Hudi 삽입, 업서트 및 삭제 작업에 큰 영향을 미칠 수 있는 성능 문제가 포함되어 있습니다. Amazon EMR 6.1.0 또는 6.2.0에서 Hudi를 사용하려는 경우 AWS 지원팀에 문의하여 패치된 Hudi RPM을 받아야 합니다.

- `spark.driver.extraJavaOptions` 및 `spark.executor.extraJavaOptions`를 사용하여 사용자 지정 가비지 수집 구성을 설정하면 가비지 수집 구성 충돌로 인해 EMR 6.1에서 드라이버 및 실행기를 시작하지 못할 수 있습니다. EMR 릴리스 6.1.0에서는 `spark.driver.defaultJavaOptions` 및 `spark.executor.defaultJavaOptions` 속성 대신, 드라이버와 실행기에 대한 사용자 지정 Spark 가비지 수집 구성을 지정해야 합니다. 자세한 내용은 [Apache Spark Runtime Environment](https://spark.apache.org/docs/latest/configuration.html#runtime-environment) 및 [Amazon EMR 6.1.0에서 Spark 가비지 수집 구성](https://docs.aws.amazon.com/emr/latest/ReleaseGuide/emr-spark-configure.html#spark-gc-config)을 참조하세요.
- Oozie에서 Pig를 사용하는 경우(또한 Hue는 Oozie 작업을 사용하여 Pig 스크립트를 실행하므로 Hue 내에서 Pig를 사용하는 경우) 네이티브 lzo 라이브러리를 로드할 수 없다는 오류가 발생합니다. 이 오류 메시지는 정보 제공용이며 Pig 실행을 차단하지 않습니다.
- Hudi 동시성 지원: 현재 Hudi는 단일 Hudi 테이블에 대한 동시 쓰기를 지원하지 않습니다. 또한 Hudi는 새 작성기를 시작하기 전에 진행 중인 작성기에서 수행하는 모든 변경 사항을 롤백합니다. 동시 쓰기는 이 메커니즘을 방해하고 경쟁 조건을 유발하여 데이터를 손상시킬 수 있습니다. 데이터 처리 워크플로의 일부로 항상 Hudi 테이블에서 단일 Hudi 작성기만 작동해야 합니다. Hudi는 동일한 Hudi 테이블에서 작동하는 여러 개의 동시 리더를 지원합니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Amazon EMR 6.1.0에는 Presto를 실행하는 클러스터에 영향을 미치는 문제가 있습니다. 오랜 기간(일 단위)이 경과한 후 클러스터에서 'su: failed to execute /bin/bash: Resource temporarily unavailable' 또는 'shell request failed on channel 0'과 같은 오류가 발생할 수 있습니다. 이 문제는 내부 Amazon EMR 프로세스(InstanceController)에서 너무 많은 경량 프로세스(LWP)를 생성하여 결국 Hadoop 사용자가 nproc 한도를 초과하기 때문에 발생합니다. 이를 통해 사용자는 추가 프로세스를 열 수 없습니다. 이 문제의 해결 방법은 EMR 6.2.0으로 업그레이드하는 것입니다.

6.1.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.14.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.14.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	2.0.0	EMR S3Select 커넥터
emrfs	2.42.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.11.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	3.2.1-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.5	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.5	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	2.2.5	HBase 명령줄 클라이언트
hbase-rest-server	2.2.5	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.5	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-2	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.2-incubating-amzn-2	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.2-incubating-amzn-2	Hudi로 Presto를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hudi-prestosql	0.5.2-incubating-amzn-2	Hudi와 함께 PrestoSQL을 실행하기 위한 번들 라이브러리.
hudi-spark	0.5.2-incubating-amzn-2	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hue-server	4.7.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.6.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.3.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.232	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.232	여러 쿼리를 실행하는 서비스.
presto-client	0.232	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
prestosql-coordinator	338	prestosql-workers의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
prestosql-worker	338	여러 쿼리를 실행하는 서비스.
prestosql-client	338	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	2.0.0	Apache Ranger 키 관리 시스템
spark-client	3.0.0-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	3.0.0-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	3.0.0-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	3.0.0-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.1.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0-preview1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.1.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-6.1.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-env	HDFS 환경의 값을 변경합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.

분류	설명
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.

분류	설명
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
prestosql-log	Presto log.properties 파일의 값을 변경합니다.
prestosql-config	Presto config.properties 파일의 값을 변경합니다.
prestosql-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
prestosql-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
prestosql-node	PrestoSQL의 node.properties 파일에서 값을 변경합니다.
prestosql-connector-blackhole	PrestoSQL의 blackhole.properties 파일에서 값을 변경합니다.

분류	설명
prestosql-connector-cassandra	PrestoSQL의 cassandra.properties 파일에서 값을 변경합니다.
prestosql-connector-hive	PrestoSQL의 hive.properties 파일에서 값을 변경합니다.
prestosql-connector-jmx	PrestoSQL의 jmx.properties 파일에서 값을 변경합니다.
prestosql-connector-kafka	PrestoSQL의 kafka.properties 파일에서 값을 변경합니다.
prestosql-connector-localfile	PrestoSQL의 localfile.properties 파일에서 값을 변경합니다.
prestosql-connector-memory	PrestoSQL의 memory.properties 파일에서 값을 변경합니다.
prestosql-connector-mongodb	PrestoSQL의 mongodb.properties 파일에서 값을 변경합니다.
prestosql-connector-mysql	PrestoSQL의 mysql.properties 파일에서 값을 변경합니다.
prestosql-connector-postgresql	PrestoSQL의 postgresql.properties 파일에서 값을 변경합니다.
prestosql-connector-raptor	PrestoSQL의 raptor.properties 파일에서 값을 변경합니다.
prestosql-connector-redis	PrestoSQL의 redis.properties 파일에서 값을 변경합니다.
prestosql-connector-redshift	PrestoSQL의 redshift.properties 파일에서 값을 변경합니다.
prestosql-connector-tpch	PrestoSQL의 tpch.properties 파일에서 값을 변경합니다.

분류	설명
prestosql-connector-tpcds	PrestoSQL의 tpcds.properties 파일에서 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.

분류	설명
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 6.0.1

6.0.1 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hudi, Hue, Livy, JupyterHubMXNet, OoziePhoenix, Presto, Spark, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS Java용 SDK	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.0.1 릴리스 정보

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.

- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.

6.0.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.14.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.5.0	EMR S3Select 커넥터
emrfs	2.39.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-0.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-0.1	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	3.2.1-amzn-0.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-0.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-0.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-0.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-0.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-0.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-0.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-0.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	3.2.1-amzn-0.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	2.2.3	HBase 명령줄 클라이언트
hbase-rest-server	2.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-0	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	3.1.2-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.0-incubating-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.0-incubating-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.0.0	Jupyter 노트북용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.230	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.230	여러 쿼리를 실행하는 서비스.
presto-client	0.230	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.4	Spark 명령줄 클라이언트.
spark-history-server	2.4.4	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.4	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.4	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0-SNAPSHOT	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.0.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-6.0.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.

분류	설명
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-env	HDFS 환경의 값을 변경합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.

분류	설명
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.

분류	설명
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.

분류	설명
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 6.0.0

6.0.0 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hudi, Hue, Livy, JupyterHubMXNet, OoziePhoenix, Presto, Spark, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
AWS Java용 SDK	1.11.828	1.11.828	1.11.711	1.11.711
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.12.10	2.12.10	2.12.10	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	2.2.5	2.2.5	2.2.3	2.2.3
HCatalog	3.1.2	3.1.2	3.1.2	3.1.2
Hadoop	3.2.1	3.2.1	3.2.1	3.2.1
Hive	3.1.2	3.1.2	3.1.2	3.1.2
Hudi	0.5.2-incubating-amzn-2	0.5.2-incubating-amzn-2	0.5.0-incubating-amzn-1	0.5.0-incubating-amzn-1

	emr-6.1.1	emr-6.1.0	emr-6.0.1	emr-6.0.0
Hue	4.7.1	4.7.1	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	-	-	-	-
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	5.0.0	5.0.0	5.0.0	5.0.0
Pig	0.17.0	0.17.0	-	-
Presto	0.232	0.232	0.230	0.230
Spark	3.0.0	3.0.0	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	-	-
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	338	338	-	-
Zeppelin	0.9.0	0.9.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

6.0.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 6.0.0에 대한 정보가 포함됩니다.

최초 릴리스 날짜: 2020년 3월 10일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.11.711
- Ganglia 버전 3.7.2
- Hadoop 버전 3.2.1
- HBase 버전 2.2.3
- HCatalog 버전 3.1.2
- Hive 버전 3.1.2
- Hudi 버전 0.5.0-incubating
- Hue 버전 4.4.0
- JupyterHub 버전 1.0.0
- Livy 버전 0.6.0
- MXNet 버전 1.5.1
- Oozie 버전 5.1.0
- Phoenix 버전 5.0.0
- Presto 버전 0.230
- Spark 버전 2.4.4
- TensorFlow 버전 1.14.0
- Zeppelin 버전 0.9.0-SNAPSHOT
- Zookeeper 버전 3.4.14
- 커넥터 및 드라이버: DynamoDB Connector 4.14.0

> **Note**
>
> Flink, Sqoop, Pig 및 Mahout는 Amazon EMR 버전 6.0.0에서만 사용할 수 있습니다.

새로운 특성

- YARN 도커 런타임 지원 - Spark 작업과 같은 YARN 애플리케이션은 이제 도커 컨테이너의 컨텍스트에서 실행될 수 있습니다. 이를 통해 Amazon EMR 클러스터에 사용자 지정 라이브러리를 설치할 필요 없이 도커 이미지의 종속성을 쉽게 정의할 수 있습니다. 자세한 내용은 Docker 통합 구성 및 Amazon EMR 6.0.0을 사용하여 Docker와 함께 Spark 애플리케이션 실행을 참조하세요.
- Hive LLAP 지원 - Hive는 이제 향상된 쿼리 성능을 위해 LLAP 실행 모드를 지원합니다. 자세한 내용은 Hive LLAP 사용을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Amazon Linux
  - Amazon Linux 2는 EMR 6.x 릴리스 시리즈의 운영 체제입니다.

- `systemd`는 Amazon Linux 1에서 사용되는 `upstart` 대신 서비스 관리에 사용됩니다.
- Java Development Kit(JDK)
  - Coretto JDK 8은 EMR 6.x 릴리스 시리즈의 기본 JDK입니다.
- Scala
  - Scala 2.12는 Apache Spark 및 Apache Livy와 함께 사용됩니다.
- Python 3
  - Python 3은 이제 EMR에서 Python의 기본 버전입니다.
- YAR의 노드 레이블
  - Amazon EMR 6.x 릴리스 시리즈부터 YARN 노드 레이블 기능은 기본적으로 비활성화되어 있습니다. 애플리케이션 마스터 프로세스는 기본적으로 코어 및 작업 노드 모두에서 실행할 수 있습니다. `yarn.node-labels.enabled` 및 `yarn.node-labels.am.default-node-label-expression` 속성을 구성하여 YARN 노드 레이블 기능을 활성화할 수 있습니다. 자세한 내용은 프라이머리, 코어 및 태스크 노드 이해를 참조하세요.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

```
LimitNOFILE=65536

LimitNPROC=65536
```

2. InstanceController 다시 시작

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- PySpark, SparkR 및 spark-shell을 포함한 Spark 대화형 셸은 도커를 추가 라이브러리와 함께 사용하는 것을 지원하지 않습니다.
- Amazon EMR 버전 6.0.0에서 Python 3을 사용하려면 `yarn.nodemanager.env-whitelist`에 `PATH`를 추가해야 합니다.
- Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용하는 경우 LLAP(Live Long and Process) 기능은 지원되지 않습니다.

- Spark 및 Docker 통합과 함께 Amazon EMR 6.0.0을 사용하는 경우 Docker 런타임으로 Spark 작업을 제출할 때 실패를 방지하기 위해 클러스터의 인스턴스를 동일한 인스턴스 유형과 동일한 크기의 EBS 볼륨으로 구성해야 합니다.
- Amazon EMR 6.0.0의 경우 Amazon S3 스토리지 모드에서 HBase는 HBASE-24286 문제의 영향을 받습니다. 기존 S3 데이터를 사용하여 클러스터를 생성할 때는 HBase 마스터를 초기화할 수 없습니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

  > **Note**
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

6.0.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	3.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.14.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.5.0	EMR S3Select 커넥터
emrfs	2.39.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	3.2.1-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	3.2.1-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	3.2.1-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	3.2.1-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	3.2.1-amzn-0	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	3.2.1-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	3.2.1-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	3.2.1-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	3.2.1-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	3.2.1-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	3.2.1-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	2.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	2.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	2.2.3	HBase 명령줄 클라이언트
hbase-rest-server	2.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	2.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	3.1.2-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	3.1.2-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	3.1.2-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	3.1.2-amzn-0	Hive 명령줄 클라이언트
hive-hbase	3.1.2-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	3.1.2-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	3.1.2-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.0-incubating-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.0-incubating-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.0.0	Jupyter 노트북용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MariaDB 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.

구성 요소	버전	설명
phoenix-library	5.0.0-HBase-2.0	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	5.0.0-HBase-2.0	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.230	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.230	여러 쿼리를 실행하는 서비스.
presto-client	0.230	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.4	Spark 명령줄 클라이언트.
spark-history-server	2.4.4	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.4	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.4	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.

구성 요소	버전	설명
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.41 이상	Apache HTTP 서버.
zeppelin-server	0.9.0-SNAPSHOT	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

6.0.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-6.0.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-env	HDFS 환경의 값을 변경합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive	Apache Hive에 대한 Amazon EMR 큐레이트 설정입니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.

분류	설명
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.

분류	설명
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 5.x 릴리스 버전

이 섹션에는 각 Amazon EMR 5.x 릴리스 버전에 제공되는 애플리케이션 버전, 릴리스 정보, 구성 요소 버전 및 구성 분류가 포함되어 있습니다.

클러스터를 시작할 때 Amazon EMR의 여러 릴리스 중에서 선택할 수 있습니다. 이를 통해 호환성 요구 사항에 맞는 애플리케이션 버전을 테스트하고 사용할 수 있습니다. 릴리스 레이블과 함께 릴리스 번호를 지정합니다. 릴리스 레이블은 `emr-x.x.x` 형식입니다. 예를 들어 `emr-7.9.0`입니다.

새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

모든 Amazon EMR 5.x 릴리스에서 애플리케이션 버전의 전체 테이블은 Amazon EMR 5.x 릴리스의 애플리케이션 버전 섹션을 참조하세요.

주제

- Amazon EMR 릴리스 5.5.2
- Amazon EMR 릴리스 5.5.1
- Amazon EMR 릴리스 5.5.0
- Amazon EMR 릴리스 5.4.1
- Amazon EMR 릴리스 5.4.0
- Amazon EMR 릴리스 5.3.2
- Amazon EMR 릴리스 5.3.1
- Amazon EMR 릴리스 5.3.0
- Amazon EMR 릴리스 5.2.3
- Amazon EMR 릴리스 5.2.2
- Amazon EMR 릴리스 5.2.1
- Amazon EMR 릴리스 5.2.0
- Amazon EMR 릴리스 5.1.1
- Amazon EMR 릴리스 5.1.0
- Amazon EMR 릴리스 5.0.3
- Amazon EMR 릴리스 5.0.2
- Amazon EMR 릴리스 5.0.1
- Amazon EMR 릴리스 5.0.0

Amazon EMR 5.x 릴리스의 애플리케이션 버전

각 Amazon EMR 5.x 릴리스에서 사용할 수 있는 애플리케이션 버전을 표시한 전체 테이블을 보려면 브라우저에서 Amazon EMR 5.x 릴리스의 애플리케이션 버전을 엽니다.

Amazon EMR 릴리스 5.36.2

5.36.2 애플리케이션 버전

이 릴리스에는 Delta, , Flink, Ganglia, HBase, HCatalog, Hive, Hadoop, Hudi, HueIceberg, JupyterEnterpriseGateway, JupyterHub, Livy, , MXNet, Mahout, Oozie, Phoenix, PrestoSpark, PigSqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
AWS Java용 SDK	206년 1월 12일	206년 1월 12일	206년 1월 12일	1.12.159
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.14.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.9
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.9
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2
Hue	4.10.0	4.10.0	4.10.0	4.10.0
Iceberg	-	-	-	-

	emr-5.36.2	emr-5.36.1	emr-5.36.0	emr-5.35.0
JupyterEnterpriseGateway	2.6.0	2.6.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.267	0.267	0.266
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.36.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.36.2에 대한 정보가 포함됩니다. 변경 사항은 5.36.1에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스에서는 Amazon EMR에서 클러스터의 HDFS 복제 인수 설정 미만으로 코어 노드를 스케일 다운하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이 개선 사항은 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 이 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항은 크기 조정 작업의 한 번 실패로 인해 클러스터 크기 조정이 무기한 실행될 위험을 줄여줍니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.
- Amazon EMR이 코어 노드를 정상적으로 해제하는 동안 클러스터 스케일 다운 작업이 중단되고 완전히 해제되기 전에 비정상 상태로 전환되는 문제를 수정합니다.
- Amazon EMR이 단일 노드를 재시작할 때 여러 프라이머리 노드가 있는 고가용성 클러스터의 노드 안정성을 개선합니다.
- Amazon EC2에서 실행되는 Amazon EMR을 사용하여 로그 관리를 최적화합니다. 따라서 클러스터 로그의 스토리지 비용이 조금 절감될 수 있습니다.
- 로그 파일이 범위를 초과하여 클러스터 작업을 방해하는 시나리오를 최소화하기 위해 프라이머리 노드에서 유지되는 ZooKeeper 트랜잭션 로그 파일의 관리를 개선합니다.
- Yarn ResourceManager와 통신할 수 없어 여러 프라이머리 노드가 있는 고가용성 클러스터에서 실패 원인이 될 수 있는 드문 버그를 수정합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오),

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240503.0	4.14.343	xxxxxx, 2024	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

5.36.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.16.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.5.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.21.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.7.0	EMR S3Select 커넥터
emrfs	2.51.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.14.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.14.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.10.1-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.1-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.9-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.9-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.9-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.9-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.9-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.9-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.9-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.10.1-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.10.1-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.10.1-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.13.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68	MySQL 데이터베이스 서버.
nvidia-cuda	11.0.194	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.267-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.267-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.267-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.

구성 요소	버전	설명
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.8-amzn-2	Spark 명령줄 클라이언트.
spark-history-server	2.4.8-amzn-2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.8-amzn-2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.8-amzn-2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.36.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.36.2 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

분류	설명	재구성 작업
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.

분류	설명	재구성 작업
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2 .properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.

분류	설명	재구성 작업
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.

분류	설명	재구성 작업
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

5.36.2 변경 로그

5.36.2 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2024-05-28	배포 완료	모든 지원되는 리전에 Amazon EMR 5.36.2 전체 배포됨
2024-05-28	문서 게시	Amazon EMR 5.36.2 릴리스 정보 최초 게시됨
2024-05-20	초기 릴리스	제한된 상용 리전에 Amazon EMR 5.36.2 최초 배포됨

Amazon EMR 릴리스 5.36.1

5.36.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Mahout, Oozie, Phoenix, Presto, PigSqoop, SparkTensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
AWS Java용 SDK	206년 1월 12일	206년 1월 12일	1.12.159	1.11.970

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.14.2	1.13.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.9	2.3.8
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.9	2.3.8
Hudi	0.10.1-amzn-1	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0
Hue	4.10.0	4.10.0	4.10.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.6.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.4.1
Livy	0.7.1	0.7.1	0.7.1	0.7.1
MXNet	1.8.0	1.8.0	1.8.0	1.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.1

	emr-5.36.1	emr-5.36.0	emr-5.35.0	emr-5.34.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.267	0.266	0.261
Spark	2.4.8	2.4.8	2.4.8	2.4.8
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.10.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.36.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.36.1에 대한 정보가 포함됩니다. 변경 사항은 5.36.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 릴리스 5.36.1은 클러스터 스케일 다운 중에 Amazon S3에 로그를 아카이브하기 위한 지원을 추가합니다. 이전 5.x 릴리스에서는 클러스터 종료 중에만 Amazon S3에 로그 파일을 아카이브할 수 있었습니다. 이 개선으로 클러스터에서 생성된 로그 파일은 노드가 종료된 후에도 Amazon S3에서 계속 유지됩니다. 자세한 내용은 클러스터 로깅 및 디버깅 구성을 참조하세요.
- 5.36.1 릴리스에서는 MR 클러스터의 추가 로그 폴더를 모니터링하도록 클러스터 내 로그 관리 대몬(daemon)이 개선되었습니다. 이 개선 사항은 디스크 과다 사용 시나리오를 최소화합니다.
- 5.36.1 릴리스에서는 클러스터 내 로그 관리 대몬(daemon)이 중지되면 해당 대몬(daemon)을 자동으로 다시 시작합니다. 이 개선 사항은 디스크 과다 사용으로 인해 노드가 비정상으로 보이는 위험을 줄입니다.

- 5.36.1 릴리스에서는 프라이머리 노드의 Amazon EMR 대몬(daemon)이 클러스터의 종료된 인스턴스에 대한 오래된 메타데이터를 유지하는 문제를 수정했습니다. 오래된 데이터를 유지 관리하면 클러스터 내 CPU 및 메모리 사용량이 제한 없이 증가하여 결국 클러스터 장애로 이어질 수 있습니다.
- 여러 프라이머리 노드에서 시작되는 클러스터와 관련하여 5.36.1 릴리스에서는 프라이머리 노드 중 하나에서 발생한 Amazon EC2 하드웨어 장애로 인해 두 번째 프라이머리 노드에서 장애가 발생하여 클러스터가 불안정해질 수 있는 문제를 수정했습니다.
- 전송 중 암호화가 구성된 클러스터와 관련하여 이제 Managed Scaling에서 Spark 셔플 데이터를 인식합니다. Spark 셔플 데이터는 Spark에서 특정 작업을 수행하기 위해 여러 파티션에서 재분배하는 데이터를 말합니다. 스케일 다운 중에 Managed Scaling은 셔플 데이터가 있는 인스턴스를 무시합니다. 이를 통해 요금 및 성능 면에서 비용이 많이 드는 작업 재시도와 재계산을 방지할 수 있습니다. 셔플 작업에 대한 자세한 내용은 Spark Programming Guide를 참조하세요.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

5.36.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.16.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.5.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.21.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.7.0	EMR S3Select 커넥터
emrfs	2.51.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.14.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.14.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.10.1-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.10.1-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.9-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.9-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.9-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.9-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.9-amzn-2	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.3.9-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.9-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.10.1-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.10.1-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.10.1-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.13.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68	MySQL 데이터베이스 서버.

구성 요소	버전	설명
nvidia-cuda	11.0.194	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.267-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.267-amzn-1	여러 쿼리를 실행하는 서비스.
presto-client	0.267-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.8-amzn-2	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.4.8-amzn-2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.8-amzn-2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.8-amzn-2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.36.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 [실행 중인 클러스터에서 인스턴스 그룹 재구성](#) 단원을 참조하십시오.

emr-5.36.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode,

분류	설명	재구성 작업
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

분류	설명	재구성 작업
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.

분류	설명	재구성 작업
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2 .properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
hudi-defaults	Hudi의 hudi-defaults.conf 파일에서 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.

분류	설명	재구성 작업
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.

분류	설명	재구성 작업
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

5.36.1 변경 로그

5.36.1 릴리스 및 릴리스 정보에 대한 변경 로그

날짜	이벤트	설명
2023-07-26	업데이트	새 OS 릴리스 레이블 `2.0.20230612.0` 및 `2.0.20230628.0` .
2023-05-25	배포 완료	모든 지원되는 리전에 Amazon EMR 5.36.1 전체 배포됨
2023-05-09	문서 게시	Amazon EMR 5.36.1 릴리스 정보 최초 게시됨
2023-05-04	초기 릴리스	제한된 상용 리전에 Amazon EMR 5.36.1 최초 배포됨

Amazon EMR 릴리스 5.36.0

5.36.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hudi, Hive, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Mahout, Oozie, Phoenix, Presto, PigSqoop, SparkTensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
AWS Java용 SDK	206년 1월 12일	1.12.159	1.11.970	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.14.2	1.13.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.9	2.3.9	2.3.8	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.9	2.3.8	2.3.7
Hudi	0.10.1-amzn-1	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1
Hue	4.10.0	4.10.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.4.1	1.2.2
Livy	0.7.1	0.7.1	0.7.1	0.7.0
MXNet	1.8.0	1.8.0	1.8.0	1.7.0

	emr-5.36.0	emr-5.35.0	emr-5.34.0	emr-5.33.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.1	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.267	0.266	0.261	0.245.1
Spark	2.4.8	2.4.8	2.4.8	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.10.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.36.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.36.0에 대한 정보가 포함됩니다. 변경 사항은 5.35.0에 관련됩니다.

최초 릴리스 날짜: 2022년 6월 15일

새로운 기능

- Amazon EMR 릴리스 5.36.0에서는 Apache Ranger 지원 클러스터에서 Apache Spark를 사용하는 데이터 정의 언어(DDL)에 대한 지원을 추가합니다. 이를 통해 Apache Ranger를 사용하여 Amazon EMR 클러스터에서 데이터베이스 및 테이블을 생성, 변경 및 삭제하는 등의 작업에 대한 액세스를 관리할 수 있습니다.

- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

> **Note**
>
> 이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, `ListReleaseLabels` API 또는 `list-release-labels` CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20230404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230119.1	4.14.301	2023년 2월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221 03.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2022004.0	4.14.294	2022년 11월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220912.1	4.14.291	2022년 10월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20220719.0	4.14.287	2022년 8월 10일	미국 서부(캘리포니아 북부), 유럽(파리), 유럽(스톡홀름), 유럽(프랑크푸르트), 아시아 태평양(뭄바이), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2022426.0	4.14.281	2022년 6월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 5.36.0 업그레이드는 이제 aws-sdk 1.12.206, Hadoop 2.10.1-amzn-4, Hive 2.3.9-amzn-2, Hudi 0.10.1-amzn-1, Spark 2.4.8-amzn-2, Presto 0.267-amzn-1, Amazon Glue connector 1.18.0, EMRFS 2.51.0을 지원합니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하

는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.

- 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
- 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

5.36.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 myapp-component라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 2.2-amzn-2로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.16.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.5.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.21.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.7.0	EMR S3Select 커넥터
emrfs	2.51.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.14.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
flink-jobmanager-config	1.14.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.10.1-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.10.1-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.1-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.9-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	2.3.9-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.9-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.9-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.9-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.9-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.9-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.10.1-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.10.1-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.10.1-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버

구성 요소	버전	설명
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.13.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68	MySQL 데이터베이스 서버.
nvidia-cuda	11.0.194	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.267-amzn-1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.267-amzn-1	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
presto-client	0.267-amzn-1	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.8-amzn-2	Spark 명령줄 클라이언트.
spark-history-server	2.4.8-amzn-2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.8-amzn-2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.8-amzn-2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.36.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.36.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and

분류	설명	재구성 작업
		Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.

분류	설명	재구성 작업
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.

분류	설명	재구성 작업
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.

분류	설명	재구성 작업
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.

분류	설명	재구성 작업
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.

분류	설명	재구성 작업
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 5.35.0

5.35.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hudi, Hive, Hue, IcebergJupyterEnterpriseGateway, JupyterHub, Livy, MXNet, , Mahout, Oozie, Phoenix, Presto, PigSqoop, SparkTensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
AWS Java용 SDK	1.12.159	1.11.970	1.11.970	1.11.970
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.14.2	1.13.1	1.12.1	1.12.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
HCatalog	2.3.9	2.3.8	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.9	2.3.8	2.3.7	2.3.7
Hudi	0.9.0-amzn-2	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1
Hue	4.10.0	4.9.0	4.9.0	4.9.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.4.1	1.2.2	1.2.2
Livy	0.7.1	0.7.1	0.7.0	0.7.0
MXNet	1.8.0	1.8.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.1	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.266	0.261	0.245.1	0.245.1
Spark	2.4.8	2.4.8	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.4.1	2.4.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.35.0	emr-5.34.0	emr-5.33.1	emr-5.33.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.10.0	0.9.0	0.9.0
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.35.0 릴리스 정보

Amazon EMR 릴리스 5.35.0의 릴리스 정보입니다.

다음 릴리스 정보에는 Amazon EMR 릴리스 5.35.0에 대한 정보가 포함됩니다. 변경 사항은 5.34.0에 관련됩니다.

최초 릴리스 날짜: 2022년 3월 30일

새로운 기능

- Log4j 1.x 및 Log4j 2.x를 사용하는 Amazon EMR 릴리스 5.35 애플리케이션은 각각 Log4j 1.2.17 이상 및 Log4j 2.17.1 이상을 사용하도록 업그레이드되었으며, 이전 릴리스의 CVE 문제를 완화하기 위해 부트스트랩 작업을 사용할 필요가 없습니다. CVE-2021-44228 완화를 위한 접근 방법을(를) 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

Flink 변경 사항

유형 변경	설명
업그레이드	• Flink 버전을 1.14.2로 업데이트합니다. • log4j를 2.17.1로 업그레이드했습니다.

Hadoop 변경 사항

유형 변경	설명
EMR 5.34.0 이후 Hadoop 오픈 소스 백포트	• YARN-10438: ClientRMService#getContainerReport()에서 null인 containerId 처리 • YARN-7266: 타임라인 서버 이벤트 핸들러 스레드 잠김 • YARN-10438: RollingLevelDb 파일이 손상되거나 누락된 경우 ATS 1.5가 시작되지 않음 • HADOOP-13500: 구성 속성 객체의 반복 동기화 • YARN-10651: AbstractYarnScheduler.updateNodeResource()에서 CapacityScheduler가 NPE와 충돌함 • HDFS-12221: XmlEditsVisitor에서 xerces 교체 • HDFS-16410: OfflineEditsXmlLoader에서 안전하지 않은 Xml 구문 분석
Hadoop 변경 사항 및 수정 사항	• KMS 및 HttpFS에서 사용되는 Tomcat이 8.5.75로 업그레이드되었습니다. • FileSystemOptimizedCommitterV2에서 커미터를 생성하는 동안 정의된 commitJob 출력 경로에 성공 마커가 기록되었습니다. CommitJob과 작업 수준 출력 경로가 다를 수 있으므로 매니페스트 파일에 정의된 경로를 사용하도록 경로가 수정되었습니다. Hive 작업의 경우 동적 파티션 또는 UNION ALL과 같은 작업을 수행할 때 성공 마커가 올바르게 기록됩니다.

Hive 변경 사항

유형 변경	설명
Hive는 이러한 JIRA 수정 사항을 포함하여 오픈 소스 릴리스 2.3.9로 업그레이드되었습니다.	• HIVE-17155: HiveConf.java 내 findConfFile()의 conf 경로에 몇 가지 문제가 있음 • HIVE-24797: Avro 스키마를 구문 분석할 때 기본값 검증 비활성화 • HIVE-21563: registerAllFunctionsOnce를 비활성화하여 Table#getEmptyTable 성능 개선 • HIVE-18147: java.net.BindException에서 테스트에 실패할 수 있음: 주소가 이미 사용 중임 • HIVE-24608: Hive 2.3.x용 HMS 클라이언트에서 get_table로 다시 전환 • HIVE-21200: 벡터화 - parquet에 대해 java.lang.UnsupportedOperationException이 발생한 날짜 열 • HIVE-19228: commons-httpclient 3.x 사용 제거
EMR 5.34.0 이후 Hive 오픈 소스 백포트	• HIVE-19990: 조인 조건에서 간격 리터럴을 사용하는 쿼리 실패 • HIVE-25824: branch-2.3을 log4j 2.17.0으로 업그레이드 • TEZ-4062: 태스크가 완료되면 추론적 시도 예약을 중단해야 함 • TEZ-4108: 추론적 실행 레이스 조건 중에 NullPointerException 발생 • TEZ-3918: tez.task.log.level 설정이 적용되지 않음
Hive 업그레이드 및 수정 사항	• Log4j 버전을 2.17.1로 업그레이드 • ORC 버전을 1.4.3으로 업그레이드

유형 변경	설명
	• ShuffleScheduler의 페널티 스레드로 인한 교착 상태 수정됨
새로운 특성	• AM 로그에 Hive 쿼리를 인쇄하는 기능이 추가되었습니다. 이 옵션은 기본적으로 비활성화되어 있습니다. 플래그/구성: `tez.am.emr.print.hive.query.in.log` . 상태(기본값): FALSE.

Oozie 변경 사항

유형 변경	설명
EMR 5.34.0 이후 Oozie 오픈 소스 백포트	• [OOZIE-3652](https://issues.apache.org/jira/browse/OOZIE-3652): NoSuchFileException 발생 시 Oozie 시작 관리자에서 디렉터리 나열을 재시도해야 함

Pig 변경 사항

유형 변경	설명
업그레이드	• log4j를 1.2.17로 업그레이드했습니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가

이 카테고리에 속합니다. `!"#$%&'()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자](#)를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.35.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.15.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.5.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경

구성 요소	버전	설명
emr-s3-dist-cp	2.20.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.7.0	EMR S3Select 커넥터
emrfs	2.49.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.14.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.14.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리

구성 요소	버전	설명
hadoop-hdfs-namenode	2.10.1-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-3	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.10.1-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.1-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.9-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.9-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.9-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.9-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.9-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.9-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.9-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.9.0-amzn-2	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.9.0-amzn-2	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.9.0-amzn-2	Hudi로 Presto를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hue-server	4.10.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.13.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68	MySQL 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.266-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.266-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.266-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.8-amzn-1	Spark 명령줄 클라이언트.
spark-history-server	2.4.8-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.8-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.8-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.35.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.35.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.

분류	설명	재구성 작업
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

분류	설명	재구성 작업
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.

분류	설명	재구성 작업
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 5.34.0

5.34.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, Oozie, , Phoenix, Pig, Spark, PrestoTez, SqoopTensorFlow, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
AWS Java용 SDK	1.11.970	1.11.970	1.11.970	1.11.890
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
Flink	1.13.1	1.12.1	1.12.1	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.8	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.8	2.3.7	2.3.7	2.3.7
Hudi	0.9.0-amzn-0	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0
Hue	4.9.0	4.9.0	4.9.0	4.8.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.4.1	1.2.2	1.2.2	1.1.0
Livy	0.7.1	0.7.0	0.7.0	0.7.0
MXNet	1.8.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.1	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.261	0.245.1	0.245.1	0.240.1
Spark	2.4.8	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7

	emr-5.34.0	emr-5.33.1	emr-5.33.0	emr-5.32.1
TensorFlow	2.4.1	2.4.1	2.4.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.10.0	0.9.0	0.9.0	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.34.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.34.0에 대한 정보가 포함됩니다. 변경 사항은 5.33.1에 관련됩니다.

최초 릴리스 날짜: 2022년 1월 20일

릴리스 날짜: 2022년 3월 21일

새로운 기능

- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.
- [Hudi] Hudi 구성을 단순화하기 위한 개선 사항. 기본적으로 낙관적 동시성 제어를 비활성화했습니다.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- 이전에는 다중 마스터 클러스터에서 리소스 관리자를 수동으로 다시 시작하면 Zookeeper와 같은 Amazon EMR 클러스터 내 대몬(daemon)이 Zookeeper znode 파일에서 이전에 해제되거나 손실된 노드를 모두 다시 로드했습니다. 이로 인해 특정 상황에서 기본 한도가 초과되었습니다. 이제

Amazon EMR은 Zookeeper 파일에서 1시간 넘게 경과된 해제되거나 손실된 노드 레코드를 제거하고 내부 한도를 늘렸습니다.

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- Zeppelin이 버전 0.10.0으로 업그레이드되었습니다.
- Livy 수정 - 0.7.1로 업그레이드됨
- Spark 성능 개선 - EMR 5.34.0에서 특정 Spark 구성 값이 재정의되면 이기종 실행기가 비활성화됩니다.
- WebHDFS 및 HttpFS 서버는 기본적으로 비활성화됩니다. Hadoop 구성(`dfs.webhdfs.enabled`)을 사용하여 WebHDFS를 다시 활성화할 수 있습니다. HttpFS 서버는 `sudo systemctl start hadoop-httpfs`를 사용하여 시작할 수 있습니다.

알려진 문제

- Livy 사용자 가장과 함께 사용되는 Amazon EMR Notebooks 기능은 기본적으로 HttpFS 가 비활성화되었기 때문에 작동하지 않습니다. 이 경우 EMR 노트북은 Livy 가장 기능이 활성화된 클러스터에 연결할 수 없습니다. 해결 방법은 클러스터에 EMR 노트북을 연결하기 전에 `sudo systemctl start hadoop-httpfs`를 사용하여 HttpFS 서버를 시작하는 것입니다.
- Apache Hadoop HttpFS 서버는 기본적으로 비활성화되어 있기 때문에 Amazon EMR 6.4.0에서는 Hue 쿼리가 작동하지 않습니다. Amazon EMR 6.4.0에서 Hue를 사용하려면 `sudo systemctl`

`start hadoop-httpfs`를 사용하여 Amazon EMR 프라이머리 노드에서 HttpFS 서버를 수동으로 시작하거나 Amazon EMR 단계를 사용합니다.

- Livy 사용자 위장과 함께 사용되는 Amazon EMR Notebooks 기능은 기본적으로 HttpFS 가 비활성화되었기 때문에 작동하지 않습니다. 이 경우 EMR 노트북은 Livy 가장 기능이 활성화된 클러스터에 연결할 수 없습니다. 해결 방법은 클러스터에 EMR 노트북을 연결하기 전에 `sudo systemctl start hadoop-httpfs`를 사용하여 HttpFS 서버를 시작하는 것입니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.34.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.14.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.4.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.18.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.7.0	EMR S3Select 커넥터
emrfs	2.48.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.13.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.13.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-2	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.10.1-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.10.1-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.8-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.8-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.8-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.8-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.8-amzn-0	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.3.8-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.8-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.9.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.9.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.9.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.9.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.4.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.1-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.13.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.8.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68	MySQL 데이터베이스 서버.

구성 요소	버전	설명
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.1	Oozie 명령줄 클라이언트
oozie-server	5.2.1	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.261-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.261-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.261-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.8-amzn-0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.4.8-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.8-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.8-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.10.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.34.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.34.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode,

분류	설명	재구성 작업
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

분류	설명	재구성 작업
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.

분류	설명	재구성 작업
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2 .properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.

분류	설명	재구성 작업
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 5.33.1

5.33.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Presto, Pig, Spark, Sqoop, TensorFlow, Tez및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
AWS Java용 SDK	1.11.970	1.11.970	1.11.890	1.11.890
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.12.1	1.11.2	1.11.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.1
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0

	emr-5.33.1	emr-5.33.0	emr-5.32.1	emr-5.32.0
Hue	4.9.0	4.9.0	4.8.0	4.8.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	2.1.0
JupyterHub	1.2.2	1.2.2	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.7.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.245.1	0.240.1	0.240.1
Spark	2.4.7	2.4.7	2.4.7	2.4.7
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.4.1	2.3.1	2.3.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.9.0	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.33.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.33.0/5.33.1에 대한 정보가 포함됩니다. 변경 사항은 5.32.0에 관련됩니다.

최초 릴리스 날짜: 2021년 4월 19일

최종 업데이트 날짜: 2021년 8월 9일

업그레이드

- Amazon Glue 커넥터를 버전 1.15.0으로 업그레이드함
- 버전 1.11.970 AWS SDK for Java 으로 업그레이드됨
- EmrFS를 버전 2.46.0으로 업그레이드함
- EMR Goodies를 버전 2.14.0으로 업그레이드함
- EMR Record Server를 버전 1.9.0으로 업그레이드함
- EMR S3 Dist CP를 버전 2.18.0으로 업그레이드함
- EMR Secret Agent를 버전 1.8.0으로 업그레이드함
- Flink를 버전 1.12.1로 업그레이드함
- Hadoop을 버전 2.10.1-amzn-1로 업그레이드함
- Hive를 버전 2.3.7-amzn-4로 업그레이드함
- Hudi를 버전 0.7.0으로 업그레이드함
- Hue를 버전 4.9.0으로 업그레이드함
- OpenCV를 버전 4.5.0으로 업그레이드함
- Presto를 버전 0.245.1-amzn-0으로 업그레이드함
- R을 버전 4.0.2로 업그레이드함
- Spark를 버전 2.4.7-amzn-1로 업그레이드함
- TensorFlow를 버전 2.4.1로 업그레이드함
- Zeppelin을 버전 0.9.0으로 업그레이드함

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Apache YARN 타임라인 서버 버전 1 및 1.5 성능 문제를 해결하도록 클러스터 구성

  Apache YARN 타임라인 서버 버전 1 및 1.5는 매우 자주 사용되는 대규모 EMR 클러스터, 특히 Amazon EMR의 기본 설정인 `yarn.resourcemanager.system-metrics-publisher.enabled=true`에서 성능 문제를 일으킬 수 있습니다. 오픈 소스 YARN 타임라인 서버 v2는 YARN 타임라인 서버 확장성과 관련된 성능 문제를 해결했습니다.

  이 문제의 다른 해결 방법은 다음과 같습니다.

  - yarn-site.xml 파일에서 yarn.resourcemanager.system-metrics-publisher.enabled=false를 구성합니다.
  - 아래 설명에 따라 클러스터를 생성할 때 이 문제를 해결하는 수정 사항을 활성화합니다.

  다음 Amazon EMR 릴리스에는 이 YARN 타임라인 서버 성능 문제에 대한 수정 사항이 포함되어 있습니다.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

위에 지정된 Amazon EMR 릴리스에서 수정 사항을 활성화하려면 `aws emr create-cluster` 명령 파라미터(`--configurations file://./configurations.json`)를 사용하여 전달되는 구성 JSON 파일에서 이러한 속성을 `true`로 설정합니다. 또는 재구성 콘솔 UI를 사용하여 수정 사항을 활성화할 수도 있습니다.

configurations.json 파일 내용 예제:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- 이제 Spark 삽입 쿼리를 위해 Hive 메타스토어에서 파티션 위치를 가져올 때 Spark 런타임이 더 빨라졌습니다.
- 업그레이드된 구성 요소 버전. 구성 요소 버전 목록은 이 안내서에서 Amazon EMR 릴리스 정보를 참조하세요.
- 각 새 클러스터에 AWS Java SDK 번들을 설치했습니다. 개별 구성 요소 jar 대신 모든 서비스 SDK와 해당 종속 항목을 포함하는 단일 jar입니다. 자세한 내용은 Java SDK Bundled Dependency를 참조하세요.
- 이전 Amazon EMR 릴리스에서 Managed Scaling 문제를 수정하고 애플리케이션 장애 비율을 크게 줄일 수 있도록 개선했습니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.

새로운 특성

- Amazon EMR은 공유 데이터 레이크에 대한 액세스를 쉽게 관리할 수 있는 Amazon S3의 기능인 Amazon S3 액세스 포인트를 지원합니다. Amazon S3 액세스 포인트 별칭을 사용하면 Amazon EMR에서 대규모 데이터 액세스를 간소화할 수 있습니다. Amazon EMR을 사용할 수 있는 모든 리전에서 추가 비용 없이 모든 버전의 Amazon EMR에서 Amazon S3 액세스 포인트를 사용할 수 있습니다. AWS Amazon S3 액세스 포인트 및 액세스 포인트 별칭에 대한 자세한 내용은 Amazon S3 사용 설명서에서 액세스 포인트에 버킷 스타일 별칭 사용을 참조하세요.
- Amazon EMR-5.33에서는 새로운 Amazon EC2 인스턴스 유형(c5a, c5ad, c6gn, c6gd, m6gd, d3, d3en, m5zn, r5b, r6gd)을 지원합니다. 지원되는 인스턴스 유형을 참조하세요.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

     ```
     LimitNOFILE=65536
     ```

     ```
     LimitNPROC=65536
     ```

  2. InstanceController 다시 시작

```
$ sudo systemctl daemon-reload

$ sudo systemctl restart instance-controller
```

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Amazon EMR 6.3.0 및 6.2.0 프라이빗 서브넷 클러스터의 경우 Ganglia 웹 UI에 액세스할 수 없습니다. '액세스 거부됨(403)' 오류가 발생합니다. Spark, Hue, JupyterHub, Zeppelin, Livy, Tez와 같은 다른 웹 UI는 정상적으로 작동합니다. 퍼블릭 서브넷 클러스터에서 Ganglia 웹 UI 액세스도 정상적으로 작동합니다. 이 문제를 해결하려면 `sudo systemctl restart httpd`를 사용하여 프라이머리 노드에서 httpd 서비스를 다시 시작합니다. 이 문제는 Amazon EMR 6.4.0에서 수정되었습니다.

- **Important**

  Amazon Linux 또는 Amazon Linux 2 Amazon Machine Image(AMI)를 실행하는 EMR 클러스터는 기본 Amazon Linux 동작을 사용하며, 재부팅이 필요한 중요한 핵심 커널 업데이트를 자동으로 다운로드하여 설치하지는 않습니다. 이는 기본 Amazon Linux AMI를 실행하는 다른 Amazon EC2 인스턴스와 동일한 동작입니다. Amazon EMR 릴리스가 출시된 후 재부팅

이 필요한 새로운 Amazon Linux 소프트웨어 업데이트(예: 커널, NVIDIA, CUDA 업데이트)를 사용할 수 있게 되면 기본 AMI를 실행하는 EMR 클러스터 인스턴스는 해당 업데이트를 자동으로 다운로드하여 설치하지 않습니다. 커널 업데이트를 받으려면 최신 Amazon Linux AMI를 사용하도록 Amazon EMR AMI를 사용자 지정할 수 있습니다.

- AWS Ranger 통합 옵션을 지정하는 보안 구성을 생성하는 콘솔 지원은 현재 GovCloud 리전에서 지원되지 않습니다. 보안 구성은 CLI를 사용하여 수행할 수 있습니다. Amazon EMR 관리 안내서에서 EMR 보안 구성 생성을 참조하세요.
- 범위 관리형 정책: AWS 모범 사례에 맞게 Amazon EMR은 더 이상 사용되지 않을 정책을 대체하기 위해 v2 EMR 범위 기본 관리형 정책을 도입했습니다. Amazon EMR 관리형 정책을 참조하세요.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&'()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.33.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.14.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.2.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.18.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.6.0	EMR S3Select 커넥터
emrfs	2.46.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.12.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.12.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-1.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-1.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-1.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-1.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-1.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-1.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-1.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.10.1-amzn-1.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-1.1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.10.1-amzn-1.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-1.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.7-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.7-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.7-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.7-amzn-4	Hive 명령줄 클라이언트
hive-hbase	2.3.7-amzn-4	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.3.7-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.7-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.7.0-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.7.0-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.7.0-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.9.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.2.2	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MySQL 데이터베이스 서버.

구성 요소	버전	설명
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.245.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.245.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.245.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.7-amzn-1.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.4.7-amzn-1.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.7-amzn-1.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.7-amzn-1.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.9.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.33.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.33.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode,

분류	설명	재구성 작업
		SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.prop erties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

분류	설명	재구성 작업
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.

분류	설명	재구성 작업
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2 .properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.

분류	설명	재구성 작업
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 5.33.0

5.33.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHub, JupyterEnterpriseGatewayLivy, MXNet, Mahout, Oozie, , Phoenix, Pig, Spark, PrestoSqoop, TensorFlow, Tez및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
AWS Java용 SDK	1.11.970	1.11.890	1.11.890	1.11.852
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.12.1	1.11.2	1.11.2	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.1	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.7.0-amzn-1	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0

	emr-5.33.0	emr-5.32.1	emr-5.32.0	emr-5.31.1
Hue	4.9.0	4.8.0	4.8.0	4.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	2.1.0	2.1.0	2.1.0	-
JupyterHub	1.2.2	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.7.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.245.1	0.240.1	0.240.1	0.238.3
Spark	2.4.7	2.4.7	2.4.7	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.4.1	2.3.1	2.3.1	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.9.0	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.33.0 릴리스 정보

5.33.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.14.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.2.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.18.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.6.0	EMR S3Select 커넥터

구성 요소	버전	설명
emrfs	2.46.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.12.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.12.1	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	2.10.1-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.10.1-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.1-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hcatalog-client	2.3.7-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.7-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.7-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.7-amzn-4	Hive 명령줄 클라이언트
hive-hbase	2.3.7-amzn-4	Hive-hbase 클라이언트
hive-metastore-server	2.3.7-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.7-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.7.0-amzn-1	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.7.0-amzn-1	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.7.0-amzn-1	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.9.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
jupyterhub	1.2.2	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68	MySQL 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.5.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.245.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
presto-worker	0.245.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.245.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	4.0.2	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.7-amzn-1	Spark 명령줄 클라이언트.
spark-history-server	2.4.7-amzn-1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.7-amzn-1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.7-amzn-1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.4.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.9.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.33.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.33.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts Hadoop KMS, Ranger KMS, HiveServe r2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServ er, and TimelineServer. Ad ditionally restarts HBaseRegi onserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore , Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-session	Kubernetes 및 Yarn 세션에 대한 Flink log4j-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.

분류	설명	재구성 작업
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.

분류	설명	재구성 작업
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.

분류	설명	재구성 작업
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.

분류	설명	재구성 작업
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.

분류	설명	재구성 작업
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.

분류	설명	재구성 작업
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.

분류	설명	재구성 작업
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zeppelin-site	zeppelin-site.xml에서 구성 설정을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.

분류	설명	재구성 작업
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 5.32.1

5.32.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hudi, Hive, Hue, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, , Oozie, Phoenix, Pig, Spark, PrestoSqoop, TensorFlow, Tez및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
AWS Java용 SDK	1.11.890	1.11.890	1.11.852	1.11.852
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
Flink	1.11.2	1.11.2	1.11.0	1.11.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.7
Hadoop	2.10.1	2.10.1	2.10.0	2.10.0
Hive	2.3.7	2.3.7	2.3.7	2.3.7
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0
Hue	4.8.0	4.8.0	4.7.1	4.7.1
Iceberg	-	-	-	-
JupyterEn terpriseGateway	2.1.0	2.1.0	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.7.0	1.6.0	1.6.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.240.1	0.240.1	0.238.3	0.238.3
Spark	2.4.7	2.4.7	2.4.6	2.4.6
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7

	emr-5.32.1	emr-5.32.0	emr-5.31.1	emr-5.31.0
TensorFlow	2.3.1	2.3.1	2.1.0	2.1.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.32.1 릴리스 정보

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.

- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.32.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.13.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.1.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.17.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.6.0	EMR S3Select 커넥터
emrfs	2.45.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.11.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.11.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-0.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-0.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-0.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-0.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-0.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-0.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-0.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.10.1-amzn-0.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-0.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.1-amzn-0.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-0.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.7-amzn-3	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.7-amzn-3	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.7-amzn-3	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.7-amzn-3	Hive 명령줄 클라이언트
hive-hbase	2.3.7-amzn-3	Hive-hbase 클라이언트
hive-metastore-server	2.3.7-amzn-3	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.7-amzn-3	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.6.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.6.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.6.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.8.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68 이상	MySQL 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.240.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.240.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.240.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.

구성 요소	버전	설명
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.7-amzn-0.1	Spark 명령줄 클라이언트.
spark-history-server	2.4.7-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.7-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.7-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.3.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.32.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.32.1 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

분류	설명	재구성 작업
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.

분류	설명	재구성 작업
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2 .properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.

분류	설명	재구성 작업
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 5.32.0

5.32.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterEnterpriseGatewayJupyterHub, Livy, MXNet, Mahout, Oozie, Phoenix, Presto, Pig, Spark, Tez, SqoopTensorFlow, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
AWS Java용 SDK	1.11.890	1.11.852	1.11.852	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.2	1.11.0	1.11.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.7	2.3.6
Hadoop	2.10.1	2.10.0	2.10.0	2.8.5
Hive	2.3.7	2.3.7	2.3.7	2.3.6
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating
Hue	4.8.0	4.7.1	4.7.1	4.6.0
Iceberg	-	-	-	-

	emr-5.32.0	emr-5.31.1	emr-5.31.0	emr-5.30.2
JupyterEnterpriseGateway	2.1.0	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.7.0	1.6.0	1.6.0	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.240.1	0.238.3	0.238.3	0.232
Spark	2.4.7	2.4.6	2.4.6	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.3.1	2.1.0	2.1.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.32.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.32.0에 대한 정보가 포함됩니다. 변경 사항은 5.31.0에 관련됩니다.

최초 릴리스 날짜: 2021년 1월 8일

업그레이드

- Amazon Glue 커넥터를 버전 1.14.0으로 업그레이드함
- Amazon SageMaker Spark SDK를 버전 1.4.1로 업그레이드함
- 버전 1.11.890 AWS SDK for Java 으로 업그레이드됨
- EMR DynamoDB Connector 버전 4.16.0을 업그레이드함
- EMRFS를 버전 2.45.0으로 업그레이드함
- EMR Log Analytics Metric을 버전 1.18.0으로 업그레이드함
- EMR MetricsAndEventsApiGateway 클라이언트를 버전 1.5.0으로 업그레이드함
- EMR Record Server를 버전 1.8.0으로 업그레이드함
- EMR S3 Dist CP를 버전 2.17.0으로 업그레이드함
- EMR Secret Agent를 버전 1.7.0으로 업그레이드함
- Flink를 버전 1.11.2로 업그레이드함
- Hadoop을 버전 2.10.1-amzn-0으로 업그레이드함
- Hive를 버전 2.3.7-amzn-3으로 업그레이드함
- Hue를 버전 4.8.0으로 업그레이드함
- Mxnet을 버전 1.7.0으로 업그레이드함
- OpenCV를 버전 4.4.0으로 업그레이드함
- Presto를 버전 0.240.1-amzn-0으로 업그레이드함
- Spark를 버전 2.4.7-amzn-0으로 업그레이드함
- TensorFlow를 버전 2.3.1로 업그레이드함

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.

- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- 업그레이드된 구성 요소 버전.
- 구성 요소 버전 목록은 이 안내서에서 Amazon EMR 릴리스 정보를 참조하세요.

새로운 특성

- Amazon EMR 5.32.0 및 6.5.0부터 Apache Spark의 동적 실행기 크기 조정이 기본적으로 활성화됩니다. 이 기능을 켜거나 끄기 위해 `spark.yarn.heterogeneousExecutors.enabled` 구성 파라미터를 사용할 수 있습니다.
- 인스턴스 메타데이터 서비스(IMDS) V2 지원 상태: Amazon EMR 5.23.1, 5.27.1 및 5.32 이상의 구성 요소는 모든 IMDS 직접 호출에 IMDSv2를 사용합니다. 애플리케이션 코드에서 IMDS를 직접 호출하는 경우, IMDSv1 및 IMDSv2를 모두 사용하거나, 추가 보안을 위해 IMDSv2만 사용하도록 IMDS를 구성할 수 있습니다. 다른 5.x EMR 릴리스의 경우 ImDSv1을 비활성화하면 클러스터 시작 오류가 발생합니다.
- Amazon EMR 5.32.0부터 Apache Ranger와 기본적으로 통합되는 클러스터를 시작할 수 있습니다. Apache Ranger는 Hadoop 플랫폼 전반에서 포괄적인 데이터 보안을 지원, 모니터링 및 관리할 수 있는 오픈 소스 프레임워크입니다. 자세한 내용은 Apache Ranger를 참조하세요. 기본 통합을 통해 자체 Apache Ranger를 이용해 Amazon EMR에서 세분화된 데이터 액세스 제어를 적용할 수 있습니다. Amazon EMR 릴리스 안내서에서 Amazon EMR과 Apache Ranger 통합을 참조하세요.

- Amazon EMR 릴리스 5.32.0은 Amazon EMR on EKS를 지원합니다. EMR on EKS를 시작하는 방법에 대한 자세한 내용은 [Amazon EMR on EKS란 무엇인가요?](#)를 참조하세요.
- Amazon EMR 릴리스 5.32.0은 Amazon EMR Studio(평가판)를 지원합니다. EMR Studio를 시작하는 방법에 대한 자세한 내용은 [Amazon EMR Studio(평가판)](#)을 참조하세요.
- 범위 관리형 정책: AWS 모범 사례에 맞게 Amazon EMR은 더 이상 사용되지 않을 정책을 대체하기 위해 v2 EMR 범위 기본 관리형 정책을 도입했습니다. [Amazon EMR 관리형 정책](#)을 참조하세요.

알려진 문제

- Amazon EMR 6.3.0 및 6.2.0 프라이빗 서브넷 클러스터의 경우 Ganglia 웹 UI에 액세스할 수 없습니다. '액세스 거부됨(403)' 오류가 발생합니다. Spark, Hue, JupyterHub, Zeppelin, Livy, Tez와 같은 다른 웹 UI는 정상적으로 작동합니다. 퍼블릭 서브넷 클러스터에서 Ganglia 웹 UI 액세스도 정상적으로 작동합니다. 이 문제를 해결하려면 `sudo systemctl restart httpd`를 사용하여 프라이머리 노드에서 httpd 서비스를 다시 시작합니다. 이 문제는 Amazon EMR 6.4.0에서 수정되었습니다.
- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

     `LimitNOFILE=65536`

     `LimitNPROC=65536`

2. InstanceController 다시 시작

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon Linux 또는 Amazon Linux 2 Amazon Machine Image(AMI)를 실행하는 EMR 클러스터는 기본 Amazon Linux 동작을 사용하며, 재부팅이 필요한 중요한 핵심 커널 업데이트를 자동으로 다운로드하여 설치하지는 않습니다. 이는 기본 Amazon Linux AMI를 실행하는 다른 Amazon EC2 인스턴스와 동일한 동작입니다. Amazon EMR 릴리스가 출시된 후 재부팅이 필요한 새로운 Amazon Linux 소프트웨어 업데이트(예: 커널, NVIDIA, CUDA 업데이트)를 사용할 수 있게 되면 기본 AMI를 실행하는 EMR 클러스터 인스턴스는 해당 업데이트를 자동으로 다운로드하여 설치하지 않습니다. 커널 업데이트를 받으려면 최신 Amazon Linux AMI를 사용하도록 Amazon EMR AMI를 사용자 지정할 수 있습니다.

- AWS Ranger 통합 옵션을 지정하는 보안 구성을 생성하는 콘솔 지원은 현재 GovCloud 리전에서 지원되지 않습니다. 보안 구성은 CLI를 사용하여 수행할 수 있습니다. Amazon EMR 관리 안내서에서 EMR 보안 구성 생성을 참조하세요.
- Amazon EMR 5.31.0 또는 5.32.0을 사용하는 클러스터에서 AtRestEncryption 또는 HDFS 암호화를 활성화하면 Hive 쿼리에서 다음과 같은 런타임 예외가 발생합니다.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.32.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*-amzn-*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.1	Amazon SageMaker Spark SDK
emr-ddb	4.16.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.13.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.1.0	Jupyter Enterprise Gateway가 포함된 EMR 노트북용 Conda 환경
emr-s3-dist-cp	2.17.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.6.0	EMR S3Select 커넥터
emrfs	2.45.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.11.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.11.2	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.1-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.1-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.1-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.1-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.1-amzn-0	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.1-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.1-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.10.1-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.1-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.1-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.1-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.7-amzn-3	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.7-amzn-3	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.7-amzn-3	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.7-amzn-3	Hive 명령줄 클라이언트
hive-hbase	2.3.7-amzn-3	Hive-hbase 클라이언트
hive-metastore-server	2.3.7-amzn-3	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.7-amzn-3	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.6.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.6.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.6.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.8.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.7.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.68	MySQL 데이터베이스 서버.
nvidia-cuda	10.1.243	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.240.1-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.240.1-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.240.1-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.

구성 요소	버전	설명
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.7-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	2.4.7-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.7-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.7-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.3.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.32.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

재구성 작업은 실행 중인 클러스터의 인스턴스 그룹에 대한 구성을 지정할 때 수행됩니다. Amazon EMR은 사용자가 수정한 분류에 대한 재구성 작업만 시작합니다. 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 단원을 참조하십시오.

emr-5.32.0 분류

분류	설명	재구성 작업
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.	Restarts the ResourceManager service.
container-executor	하둡 YARN의 container-executor.cfg 파일의 값을 변경합니다.	Not available.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.	Not available.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Ranger KMS, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.

분류	설명	재구성 작업
docker-conf	Docker 관련 설정을 변경합니다.	Not available.
emrfs-site	EMRFS 설정을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts HBaseRegionserver, HBaseMaster, HBaseThrift, HBaseRest, HiveServer2, Hive MetaStore, Hadoop Httpfs, and MapReduce-HistoryServer.
flink-conf	flink-conf.yaml 설정을 변경합니다.	Restarts Flink history server.
flink-log4j	Flink log4j.properties 설정을 변경합니다.	Restarts Flink history server.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.	Not available.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.	Restarts Flink history server.

분류	설명	재구성 작업
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts PhoenixQueryserver, HiveServer2, Hive MetaStore, and MapReduce-HistoryServer.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Hadoop KMS, Hadoop Httpfs, and MapReduce-HistoryServer.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.	Not available.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.	Not available.

분류	설명	재구성 작업
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.	Custom EMR specific property. Sets emrfs-site and hbase-site configs. See those for their associated restarts.
hbase-env	HBase 환경의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.	Not available.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.	Restarts the HBase services RegionServer, HBaseMaster, ThriftServer, RestServer. Additionally restarts Phoenix QueryServer.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.	Should not be reconfigured.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.	Restarts the Hadoop HDFS services Namenode, SecondaryNamenode, Datanode, ZKFC, and Journalnode. Additionally restarts Hadoop Httpfs.

분류	설명	재구성 작업
hcatalog-env	HCatalog 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.	Restarts Hive HCatalog Server.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.	Restarts Hive WebHCat Server.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.	Restarts Hive WebHCat Server.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.	Not available.
hive-env	Hive 환경의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore.

분류	설명	재구성 작업
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2 .properties 파일의 값을 변경합니다.	Not available.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.	Not available.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.	Restarts HiveServer2 and HiveMetastore. Runs Hive schemaTool CLI commands to verify hive-metastore. Also restarts Oozie and Zeppelin.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.	Not available.
hue-ini	Hue ini 파일의 값을 변경합니다.	Restarts Hue. Also activates Hue config override CLI commands to pick up new configurations.
httpfs-env	HTTPFS 환경의 값을 변경합니다.	Restarts Hadoop Httpfs service.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.	Restarts Hadoop Httpfs service.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.	Not available.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.	Restarts Hadoop-KMS service.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.	Restarts Hadoop-KMS and Ranger-KMS service.
hudi-env	Hudi 환경의 값을 변경합니다.	Not available.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.	Not available.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성	Not available.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.	Not available.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.	Restarts Livy Server.
livy-env	Livy 환경의 값을 변경합니다.	Restarts Livy Server.
livy-log4j	Livy log4j.properties 설정을 변경합니다.	Restarts Livy Server.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.	Restarts Hadoop MapReduce-HistoryServer.
oozie-env	Oozie 환경의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.	Restarts Oozie.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.	Restarts Oozie.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.	Not available.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.	Not available.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.	Restarts Phoenix-QueryServer.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.	Not available.
pig-env	Pig 환경에서 값을 변경합니다.	Not available.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.	Restarts Oozie.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.	Not available.
presto-log	Presto log.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-config	Presto config.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.	Restarts Presto-Server.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.	Not available.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.	Not available.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.	Not available.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.	Restarts Presto-Server.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.	Not available.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.	Not available.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.	Not available.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.	Not available.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.	Not available.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.	Not available.

분류	설명	재구성 작업
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.	Not available.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.	Not available.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.	Not available.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site .xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.	Restarts Ranger KMS Server.
ranger-kms-log4j	Ranger KM에서 kms-log4j .properties 파일의 값을 변경합니다.	Not available.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.	Not available.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.	Restarts EMR record server.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.	Restarts EMR record server.

분류	설명	재구성 작업
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.	Restarts EMR record server.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.	This property modifies spark-defaults. See actions there.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-env	the Spark 환경의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.	Not available.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.	Restarts Spark history server and Spark thrift server.
sqoop-env	Sqoop 환경의 값을 변경합니다.	Not available.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.	Not available.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.	Not available.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.	Restarts Oozie.

분류	설명	재구성 작업
yarn-env	YARN 환경의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts MapReduce-HistoryServer.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.	Restarts the Hadoop YARN services ResourceManager, NodeManager, ProxyServer, and TimelineServer. Additionally restarts Livy Server and MapReduce-HistoryServer.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.	Restarts Zeppelin.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.	Restarts Zookeeper server.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.	Restarts Zookeeper server.

Amazon EMR 릴리스 5.31.1

5.31.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HudiHue, JupyterHubLivy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- [Amazon EMR 7.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 6.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
AWS Java용 SDK	1.11.852	1.11.852	1.11.759	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.11.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.7	2.3.6	2.3.6
Hadoop	2.10.0	2.10.0	2.8.5	2.8.5
Hive	2.3.7	2.3.7	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.7.1	4.6.0	4.6.0
Iceberg	-	-	-	-

	emr-5.31.1	emr-5.31.0	emr-5.30.2	emr-5.30.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.6.0	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.238.3	0.238.3	0.232	0.232
Spark	2.4.6	2.4.6	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	2.1.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.31.1 릴리스 정보

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.

- 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
- 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.31.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.0	Amazon SageMaker Spark SDK
emr-ddb	4.15.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.13.0	편리한 하둡 에코시스템용 추가 라이브러리

구성 요소	버전	설명
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.15.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.6.0	EMR S3Select 커넥터
emrfs	2.43.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.11.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.11.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.0-amzn-0.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.0-amzn-0.1	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.10.0-amzn-0.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.0-amzn-0.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.0-amzn-0.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.0-amzn-0.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.0-amzn-0.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.10.0-amzn-0.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.0-amzn-0.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.0-amzn-0.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.0-amzn-0.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.7-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.7-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	2.3.7-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.7-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.7-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.7-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.7-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.6.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.6.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.

구성 요소	버전	설명
hudi-presto	0.6.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.7.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.6.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.3.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.238.3-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.238.3-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.238.3-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.6-amzn-0.1	Spark 명령줄 클라이언트.
spark-history-server	2.4.6-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.6-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.6-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tensorflow	2.1.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.31.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.31.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.

분류	설명
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.

분류	설명
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.31.0

5.31.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HudiHue, JupyterHubLivy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
AWS Java용 SDK	1.11.852	1.11.759	1.11.759	1.11.759
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.7
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.11.0	1.10.0	1.10.0	1.10.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.13
HCatalog	2.3.7	2.3.6	2.3.6	2.3.6
Hadoop	2.10.0	2.8.5	2.8.5	2.8.5
Hive	2.3.7	2.3.6	2.3.6	2.3.6
Hudi	0.6.0-amzn-0	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating
Hue	4.7.1	4.6.0	4.6.0	4.6.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.1.0

	emr-5.31.0	emr-5.30.2	emr-5.30.1	emr-5.30.0
Livy	0.7.0	0.7.0	0.7.0	0.7.0
MXNet	1.6.0	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.2.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.238.3	0.232	0.232	0.232
Spark	2.4.6	2.4.5	2.4.5	2.4.5
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	2.1.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.31.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.31.0에 대한 정보가 포함됩니다. 변경 사항은 5.30.1에 관련됩니다.

최초 릴리스 날짜: 2020년 10월 9일

최종 업데이트 날짜: 2020년 10월 15일

업그레이드

- Amazon Glue 커넥터를 버전 1.13.0으로 업그레이드함
- Amazon SageMaker Spark SDK를 버전 1.4.0으로 업그레이드함
- Amazon Kinesis 커넥터를 버전 3.5.9로 업그레이드함
- 버전 1.11.852 AWS SDK for Java 로 업그레이드됨
- Bigtop-tomcat을 버전 8.5.56으로 업그레이드함
- EMR FS를 버전 2.43.0으로 업그레이드함
- EMR MetricsAndEventsApiGateway 클라이언트를 버전 1.4.0으로 업그레이드함
- EMR S3 Dist CP를 버전 2.15.0으로 업그레이드함
- EMR S3 Select를 버전 1.6.0으로 업그레이드함
- Flink를 버전 1.11.0으로 업그레이드함
- Hadoop을 버전 2.10.0으로 업그레이드함
- Hive를 버전 2.3.7로 업그레이드함
- Hudi를 버전 0.6.0으로 업그레이드함
- Hue를 버전 4.7.1로 업그레이드함
- JupyterHub가 버전 1.1.0으로 업그레이드됨
- Mxnet을 버전 1.6.0으로 업그레이드함
- OpenCV를 버전 4.3.0으로 업그레이드함
- Presto를 버전 0.238.3으로 업그레이드함
- TensorFlow를 버전 2.1.0으로 업그레이드함

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.

- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Hive 열 통계는 Amazon EMR 버전 5.31.0 이상에서 지원됩니다.
- 업그레이드된 구성 요소 버전.
- Amazon EMR 5.31.0에서 EMRFS S3EC V2 지원. S3 Java SDK 릴리스 1.11.837 이상에서 암호화 클라이언트 버전 2(S3EC V2)가 다양한 보안 개선 사항과 함께 도입되었습니다. 자세한 내용은 다음 자료를 참조하세요.
  - S3 블로그 게시물: Updates to the Amazon S3 encryption client.
  - AWS SDK for Java 개발자 안내서: 암호화 및 복호화 클라이언트를 V2로 마이그레이션합니다.
  - EMR 관리 안내서: Amazon S3 클라이언트 측 암호화

  암호화 클라이언트 V1은 이전 버전과의 호환성을 위해 SDK에서 계속 사용할 수 있습니다.

새로운 특성

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한

보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`
2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
```

```
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Amazon EMR 5.31.0에서는 Lake Formation과 통합되는 클러스터를 시작할 수 있습니다. 이 통합은 AWS Glue 데이터 카탈로그의 데이터베이스 및 테이블에 대한 세분화된 열 수준 데이터 필터링을 제공합니다. 또한 엔터프라이즈 자격 증명 시스템에서 EMR 노트북 또는 Apache Zeppelin에 연동된 Single Sign-On을 사용할 수도 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [Amazon EMR과 AWS Lake Formation통합](#)을 참조하세요.

  Lake Formation을 사용하는 Amazon EMR은 현재 미국 동부(오하이오 및 버지니아 북부), 미국 서부(캘리포니아 북부 및 오레곤), 아시아 태평양(뭄바이, 서울, 싱가포르, 시드니 및 도쿄), 캐나다(중부), 유럽(프랑크푸르트, 아일랜드, 런던, 파리 및 스톡홀름), 남아메리카(상파울루)의 16개 AWS 리전에서 사용할 수 있습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Amazon EMR 5.31.0 또는 5.32.0을 사용하는 클러스터에서 AtRestEncryption 또는 HDFS 암호화를 활성화하면 Hive 쿼리에서 다음과 같은 런타임 예외가 발생합니다.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.31.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.4.0	Amazon SageMaker Spark SDK
emr-ddb	4.15.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.13.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.15.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.6.0	EMR S3Select 커넥터
emrfs	2.43.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.11.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
flink-jobmanager-config	1.11.0	Apache Flink JobManager에 대한 EMR 노드에서 리소스를 관리합니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.10.0-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.10.0-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.10.0-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.10.0-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.10.0-amzn-0	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.10.0-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.10.0-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.10.0-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.10.0-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.10.0-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.10.0-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.7-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.7-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.7-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.7-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.7-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.7-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.7-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.6.0-amzn-0	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-spark	0.6.0-amzn-0	Hudi와 함께 Spark를 실행하기 위한 번들 라이브러리.
hudi-presto	0.6.0-amzn-0	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.7.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.6.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	4.3.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.238.3-amzn-0	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.238.3-amzn-0	여러 쿼리를 실행하는 서비스.
presto-client	0.238.3-amzn-0	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.

구성 요소	버전	설명
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.6-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	2.4.6-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.6-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.6-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	2.1.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.31.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.31.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성

분류	설명
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.

분류	설명
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.30.2

5.30.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHubLivy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
AWS Java용 SDK	1.11.759	1.11.759	1.11.759	1.11.682
Python	2.7, 3.7	2.7, 3.7	2.7, 3.7	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.10.0	1.9.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.13	1.4.10

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.6.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.1.0	1.0.0
Livy	0.7.0	0.7.0	0.7.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.2.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.232	0.227
Spark	2.4.5	2.4.5	2.4.5	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.30.2	emr-5.30.1	emr-5.30.0	emr-5.29.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.30.2 릴리스 정보

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.

- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. 공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.30.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.13.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.14.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.5.0	EMR S3Select 커넥터
emrfs	2.40.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.10.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-6.1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.8.5-amzn-6.1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-6.1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-6.1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-6.1	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-6.1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-6.1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-6.1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-6.1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-6.1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-6.1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.6-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.6-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.6-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.6-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.6-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.6-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.6-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.2-incubating	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.

구성 요소	버전	설명
hudi-presto	0.5.2-incubating	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.6.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64 이상	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.232	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.232	여러 쿼리를 실행하는 서비스.
presto-client	0.232	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.5-amzn-0.1	Spark 명령줄 클라이언트.
spark-history-server	2.4.5-amzn-0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.5-amzn-0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.5-amzn-0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.30.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.30.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.

분류	설명
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.

분류	설명
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.30.1

5.30.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHubLivy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
AWS Java용 SDK	1.11.759	1.11.759	1.11.682	1.11.659
Python	2.7, 3.7	2.7, 3.7	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.10.0	1.9.1	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.13	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.6.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.1.0	1.0.0	1.0.0

	emr-5.30.1	emr-5.30.0	emr-5.29.0	emr-5.28.1
Livy	0.7.0	0.7.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.2.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.232	0.227	0.227
Spark	2.4.5	2.4.5	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.30.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.30.1에 대한 정보가 포함됩니다. 변경 사항은 5.30.0에 관련됩니다.

최초 릴리스 날짜: 2020년 6월 30일

최종 업데이트 날짜: 2020년 8월 24일

변경 사항, 향상된 기능 및 해결된 문제

- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- 인스턴스 컨트롤러 프로세스에서 무한한 수의 프로세스를 생성하는 문제가 수정되었습니다.
- Hue에서 Hive 쿼리를 실행할 수 없어서 'database is locked' 메시지를 표시하고 쿼리가 실행되지 않는 문제가 수정되었습니다.
- EMR 클러스터에서 더 많은 작업을 동시에 실행할 수 있도록 Spark 문제가 수정되었습니다.
- Jupyter 서버에서 'too many files open error'를 발생시키는 Jupyter Notebook 문제가 수정되었습니다.
- 클러스터 시작 시간 관련 문제가 수정되었습니다.

새로운 특성

- Amazon EMR 버전 6.x 및 EMR 버전 5.30.1 이상에서 Tez UI 및 YARN 타임라인 서버의 영구 애플리케이션 인터페이스를 사용할 수 있습니다. 영구 애플리케이션 기록에 대한 원클릭 링크 액세스를 사용하면 SSH 연결을 통해 웹 프록시를 설정하지 않고도 작업 기록에 빠르게 액세스할 수 있습니다. 활성 클러스터와 종료된 클러스터에 대한 로그는 애플리케이션 종료 후 30일 동안 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [영구 애플리케이션 사용자 인터페이스 보기]를 참조하세요.
- EMR 노트북 실행 API는 스크립트 또는 명령줄을 통해 EMR 노트북을 실행하는 데 사용할 수 있습니다. AWS 콘솔 없이 EMR 노트북 실행을 시작, 중지, 나열 및 설명할 수 있으므로 EMR 노트북을 프로그래밍 방식으로 제어할 수 있습니다. 파라미터화된 노트북 셀을 사용하면 새로운 파라미터 값 세트마다 노트북의 복사본을 만들지 않고도 노트북에 다양한 파라미터 값을 전달할 수 있습니다. [EMR API 작업]을 참조하세요. 샘플 코드는 [EMR Notebooks를 프로그래밍 방식으로 실행하는 샘플 명령]을 참조하세요.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사

항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`

2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
```

```
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- EMR Notebooks

  클러스터 프라이머리 노드에 커널과 추가 Python 라이브러리를 설치할 수 있는 기능은 EMR 버전 5.30.1에서 기본적으로 비활성화됩니다. 이 기능에 대한 자세한 내용은 클러스터 프라이머리 노드에 커널 및 Python 라이브러리 설치를 참조하세요.

  이 기능을 활성화하려면 다음을 수행합니다.

  1. EMR Notebooks의 서비스 역할에 연결된 권한 정책이 다음 작업을 허용하는지 확인합니다.

     ```
     elasticmapreduce:ListSteps
     ```

     자세한 내용은 EMR 노트북 서비스 역할 단원을 참조하십시오.
  2. AWS CLI 를 사용하여 다음 예제와 같이 EMR 노트북을 설정하는 클러스터에서 단계를 실행합니다. *us-east-1*을 클러스터가 상주하는 리전으로 바꿉니다. 자세한 내용은 AWS CLI를 사용하여 클러스터에 단계 추가를 참조하세요.

     ```
     aws emr add-steps  --cluster-id MyClusterID --steps
      Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
     east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
     awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
     ```

- Managed Scaling

  Presto가 설치되지 않은 5.30.0 및 5.30.1 클러스터에서 Managed Scaling을 수행하면 애플리케이션 장애가 발생하거나 균일한 인스턴스 그룹 또는 인스턴스 플릿이 ARRESTED 상태로 유지될 수 있습니다. 특히 스케일 다운 작업 이후 바로 스케일 업 작업이 수행되는 경우가 이에 해당합니다.

  해결 방법으로 작업에 Presto가 필요하지 않더라도 Amazon EMR 릴리스 5.30.0 및 5.30.1에서 클러스터를 생성할 때 설치할 애플리케이션으로 Presto를 선택합니다.

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작

업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 /etc/hadoop.keytab에 있으며, 보안 주체는 hadoop/<hostname>@<REALM> 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 s3://bucket/table/p=a는 s3://bucket/table/p=a b의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 s3://bucket/table/p=a b에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. !"#$%&‘()*+,-와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 spark-defaults 분류에서 spark.sql.sources.fastS3PartitionDiscovery.enabled 구성을 false로 설정합니다.

5.30.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.13.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.14.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.5.0	EMR S3Select 커넥터
emrfs	2.40.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.10.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-6	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.8.5-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.6-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.6-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.6-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.6-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.6-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.6-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.6-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.2-incubating	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.2-incubating	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.6.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.

구성 요소	버전	설명
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.232	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.232	여러 쿼리를 실행하는 서비스.
presto-client	0.232	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트

구성 요소	버전	설명
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.5-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	2.4.5-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.5-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.5-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.30.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.30.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성

분류	설명
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.

분류	설명
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.30.0

5.30.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hudi, Hue, JupyterHubLivy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
AWS Java용 SDK	1.11.759	1.11.682	1.11.659	1.11.659
Python	2.7, 3.7	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.10.0	1.9.1	1.9.0	1.9.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.13	1.4.10	1.4.10	1.4.10

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
HCatalog	2.3.6	2.3.6	2.3.6	2.3.6
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.6
Hudi	0.5.2-incubating	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating
Hue	4.6.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	1.1.0	1.0.0	1.0.0	1.0.0
Livy	0.7.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.5.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.2.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.3
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.232	0.227	0.227	0.227
Spark	2.4.5	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.30.0	emr-5.29.0	emr-5.28.1	emr-5.28.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.2
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.30.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.30.0에 대한 정보가 포함됩니다. 변경 사항은 5.29.0에 관련됩니다.

최초 릴리스 날짜: 2020년 5월 13일

최종 업데이트 날짜: 2020년 6월 25일

업그레이드

- 버전 1.11.759 AWS SDK for Java 로 업그레이드됨
- Amazon SageMaker Spark SDK가 버전 1.3.0으로 업그레이드됨
- EMR 레코드 서버가 버전 1.6.0으로 업그레이드됨
- Flink가 버전 1.10.0으로 업그레이드됨
- Ganglia가 버전 3.7.2로 업그레이드됨
- HBase가 버전 1.4.13으로 업그레이드됨
- Hudi가 버전 0.5.2-incubating으로 업그레이드됨
- Hue가 버전 4.6.0으로 업그레이드됨
- JupyterHub가 버전 1.1.0으로 업그레이드됨
- Livy가 버전 0.7.0-incubating으로 업그레이드됨
- Oozie가 버전 5.2.0으로 업그레이드됨
- Presto가 버전 0.232로 업그레이드됨
- Spark가 버전 2.4.5로 업그레이드됨
- 업그레이드된 커넥터 및 드라이버: Amazon Glue 커넥터 1.12.0, Amazon Kinesis 커넥터 3.5.0, EMR DynamoDB 커넥터 4.14.0

새로운 특성

- EMR Notebooks - 5.30.0을 사용하여 생성된 EMR 클러스터와 함께 사용하면 EMR 노트북 커널이 클러스터에서 실행됩니다. 이로써 노트북 성능이 향상되며, 커널을 설치하고 사용자 지정할 수 있습니다. 클러스터 프라이머리 노드에 Python 라이브러리를 설치할 수도 있습니다. 자세한 내용은 EMR 관리 안내서에서 커널 및 라이브러리 설치 및 사용을 참조하세요.
- Managed Scaling - Amazon EMR 버전 5.30.0 이상에서는 EMR Managed Scaling을 활성화하여 워크로드에 따라 클러스터에서 인스턴스 또는 유닛 수를 자동으로 늘리거나 줄일 수 있습니다. Amazon EMR은 클러스터 지표를 지속적으로 평가하여 비용과 속도 측면에서 클러스터를 최적화하는 조정 결정을 내립니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 리소스 조정을 참조하세요.
- Amazon S3에 저장된 로그 파일 암호화 - Amazon EMR 버전 5.30.0 이상에서는 AWS KMS 고객 관리형 키를 사용하여 Amazon S3에 저장된 로그 파일을 암호화할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon S3에 저장된 로그 파일 암호화를 참조하세요.
- Amazon Linux 2 지원 - EMR 버전 5.30.0 이상에서 EMR은 Amazon Linux 2 OS를 사용합니다. 새로운 사용자 지정 Amazon Machine Image(AMI)는 Amazon Linux 2 AMI를 기반으로 해야 합니다. 자세한 내용은 사용자 지정 AMI 사용을 참조하십시오.
- Presto 단계적 자동 조정 - 5.30.0 버전을 사용하는 EMR 클러스터의 경우 Auto Scaling 제한 시간을 설정할 수 있습니다. 이 제한 시간을 통해 Presto에서 노드가 서비스 해제되기 전에 실행을 완료할 수 있습니다. 자세한 내용은 단계적 서비스 해제 기능이 있는 자동 조정 사용 단원을 참조하십시오.
- 새 할당 전략 옵션을 사용한 플릿 인스턴스 생성 - 새 할당 전략 옵션은 EMR 버전 5.12.1 이상에서 사용할 수 있습니다. 이를 통해 클러스터 프로비저닝 속도가 빨라지고 스팟 할당이 더 정확해지며 스팟 인스턴스 중단이 줄어듭니다. 기본이 아닌 EMR 서비스 역할은 업데이트해야 합니다. 인스턴스 플릿 구성을 참조하세요.
- sudo systemctl stop 및 sudo systemctl start 명령 – Amazon Linux 2 OS를 사용하는 EMR 버전 5.30.0 이상에서 EMR은 `sudo systemctl stop` 및 `sudo systemctl start` 명령을 사용하여 서비스를 다시 시작합니다. 자세한 내용은 Amazon EMR에서 서비스를 다시 시작하려면 어떻게 해야 하나요?를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- EMR 버전 5.30.0은 기본적으로 Ganglia를 설치하지 않습니다. 클러스터 생성 시 Ganglia를 설치하도록 명시적으로 선택할 수 있습니다.
- Spark 성능 최적화
- Presto 성능 최적화

- Python 3은 Amazon EMR 버전 5.30.0 이상의 기본값입니다.
- 프라이빗 서브넷의 서비스 액세스에 대한 기본 관리형 보안 그룹이 새 규칙으로 업데이트되었습니다. 서비스 액세스에 대한 사용자 지정 보안 그룹을 사용하는 경우, 기본 관리형 보안 그룹과 동일한 규칙을 포함시켜야 합니다. 자세한 내용은 [서비스 액세스에 대한 Amazon EMR 관리형 보안 그룹(프라이빗 서브넷)]을 참조하세요. Amazon EMR에 대해 사용자 지정 서비스 역할을 사용하는 경우 EMR에서 보안 그룹이 올바르게 생성되는지 여부를 검증할 수 있도록 `ec2:describeSecurityGroups`에 대한 권한을 부여해야 합니다. `EMR_DefaultRole`을 사용하는 경우 이 권한은 이미 기본 관리형 정책에 포함되어 있습니다.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

     `LimitNOFILE=65536`

     `LimitNPROC=65536`
  2. InstanceController 다시 시작

     ```
     $ sudo systemctl daemon-reload
     ```

```
$ sudo systemctl restart instance-controller
```

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Managed Scaling

  Presto가 설치되지 않은 5.30.0 및 5.30.1 클러스터에서 Managed Scaling을 수행하면 애플리케이션 장애가 발생하거나 균일한 인스턴스 그룹 또는 인스턴스 플릿이 `ARRESTED` 상태로 유지될 수 있습니다. 특히 스케일 다운 작업 이후 바로 스케일 업 작업이 수행되는 경우가 이에 해당합니다.

  해결 방법으로 작업에 Presto가 필요하지 않더라도 Amazon EMR 릴리스 5.30.0 및 5.30.1에서 클러스터를 생성할 때 설치할 애플리케이션으로 Presto를 선택합니다.

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일

다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Hue 4.6.0의 기본 데이터베이스 엔진은 SQLite이며, 이 때문에 외부 데이터베이스에서 Hue를 사용하려고 하면 문제가 발생합니다. 이 문제를 해결하려면 `hue-ini` 구성 분류에서 `engine`을 `mysql`로 설정합니다. 이 문제는 Amazon EMR 버전 5.30.1에서 수정되었습니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자]를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

5.30.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.3.0	Amazon SageMaker Spark SDK
emr-ddb	4.14.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.13.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.5.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-notebook-env	1.0.0	emr 노트북용 Conda env
emr-s3-dist-cp	2.14.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.5.0	EMR S3Select 커넥터
emrfs	2.40.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.

구성 요소	버전	설명
flink-client	1.10.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-6	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.8.5-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.13	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.13	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.13	HBase 명령줄 클라이언트
hbase-rest-server	1.4.13	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.13	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.6-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.6-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.6-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.6-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.6-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.6-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.6-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.2-incubating	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.2-incubating	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.6.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.1.0	Jupyter Notebook용 다중 사용자 서버
livy-server	0.7.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.

구성 요소	버전	설명
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mariadb-server	5.5.64	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.2.0	Oozie 명령줄 클라이언트
oozie-server	5.2.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.232	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.232	여러 쿼리를 실행하는 서비스.
presto-client	0.232	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.3	통계 컴퓨팅용 R 프로젝트

구성 요소	버전	설명
ranger-kms-server	1.2.0	Apache Ranger 키 관리 시스템
spark-client	2.4.5-amzn-0	Spark 명령줄 클라이언트.
spark-history-server	2.4.5-amzn-0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.5-amzn-0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.5-amzn-0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.30.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.30.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
hudi-env	Hudi 환경의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성

분류	설명
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.

분류	설명
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.29.0

5.29.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HudiHue, JupyterHubLivy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
AWS Java용 SDK	1.11.682	1.11.659	1.11.659	1.11.615
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.1	1.9.0	1.9.0	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
HCatalog	2.3.6	2.3.6	2.3.6	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.6	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	0.5.0-incubating	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.5.1	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.3	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.227	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.29.0	emr-5.28.1	emr-5.28.0	emr-5.27.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.2	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.29.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.29.0에 대한 정보가 포함됩니다. 변경 사항은 5.28.1에 관련됩니다.

최초 릴리스 날짜: 2020년 1월 17일

업그레이드

- 버전 1.11.682 AWS SDK for Java 로 업그레이드됨
- Hive가 버전 2.3.6으로 업그레이드됨
- Flink가 버전 1.9.1로 업그레이드됨
- EmrFS가 버전 2.38.0으로 업그레이드됨
- EMR DynamoDB Connector가 버전 4.13.0으로 업그레이드됨

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Spark 성능 최적화
- EMRFS
  - 일관된 보기를 위해 관리 안내서에서 emrfs-site.xml 기본 설정이 업데이트됩니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작

업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.29.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.13.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.12.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.13.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.4.0	EMR S3Select 커넥터
emrfs	2.38.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터
flink-client	1.9.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.8.5-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-5	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.10	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.4.10	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.10	HBase 명령줄 클라이언트
hbase-rest-server	1.4.10	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.10	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.6-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.6-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.6-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.6-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.6-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.6-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.6-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.0-incubating	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.

구성 요소	버전	설명
hudi-presto	0.5.0-incubating	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.0.0	Jupyter 노트북용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.227	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.227	여러 쿼리를 실행하는 서비스.
presto-client	0.227	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.4	Spark 명령줄 클라이언트.
spark-history-server	2.4.4	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.4	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.4	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.

구성 요소	버전	설명
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.29.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.29.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.

분류	설명
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.

분류	설명
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.

분류	설명
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.28.1

5.28.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HudiHue, JupyterHubLivy, MXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
AWS Java용 SDK	1.11.659	1.11.659	1.11.615	1.11.615
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.9.0	1.8.1	1.8.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10
HCatalog	2.3.6	2.3.6	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.6	2.3.5	2.3.5
Hudi	0.5.0-incubating	0.5.0-incubating	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	1.0.0

	emr-5.28.1	emr-5.28.0	emr-5.27.1	emr-5.27.0
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.5.1	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.3	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.227	0.224	0.224
Spark	2.4.4	2.4.4	2.4.4	2.4.4
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.14.0
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.2	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.28.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.28.1에 대한 정보가 포함됩니다. 변경 사항은 5.28.0에 관련됩니다.

최초 릴리스 날짜: 2020년 1월 10일

변경 사항, 향상된 기능 및 해결된 문제

- Spark

  - Spark 호환성 문제를 수정했습니다.
- CloudWatch 지표
  - 여러 프라이머리 노드가 있는 EMR 클러스터에 Amazon CloudWatch 지표가 게시되는 문제를 수정했습니다.
- 로그 메시지 비활성화
  - false 로그 메시지 "이전 버전(<4.5.8)의 Apache http 클라이언트를 사용..."을 비활성화했습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

  > **Note**
  >
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.28.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 emr 또는 aws로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 myapp-component라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 2.2-amzn-2로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.12.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.11.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.13.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.37.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.9.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-5	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.8.5-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.10	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.10	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.10	HBase 명령줄 클라이언트
hbase-rest-server	1.4.10	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.10	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.6-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.6-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.6-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.6-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.6-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.6-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.6-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.0-incubating	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.0-incubating	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.0.0	Jupyter 노트북용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.

구성 요소	버전	설명
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.227	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.227	여러 쿼리를 실행하는 서비스.
presto-client	0.227	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트

구성 요소	버전	설명
spark-client	2.4.4	Spark 명령줄 클라이언트.
spark-history-server	2.4.4	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.4	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.4	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.28.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 hive-site.xml과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.28.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.

분류	설명
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.28.0

5.28.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hudi, Hive, Hue, Livy, JupyterHubMXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
AWS Java용 SDK	1.11.659	1.11.615	1.11.615	1.11.595
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.9.0	1.8.1	1.8.1	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.10

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
HCatalog	2.3.6	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.6	2.3.5	2.3.5	2.3.5
Hudi	0.5.0-incubating	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	1.0.0	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.5.1	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.3	4.14.2	4.14.2	4.14.2
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.227	0.224	0.224	0.220
Spark	2.4.4	2.4.4	2.4.4	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.14.0	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2

	emr-5.28.0	emr-5.27.1	emr-5.27.0	emr-5.26.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.2	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.28.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.28.0에 대한 정보가 포함됩니다. 변경 사항은 5.27.0에 관련됩니다.

최초 릴리스 날짜: 2019년 11월 12일

업그레이드

- Flink가 버전 1.9.0으로 업그레이드됨
- Hive가 버전 2.3.6으로 업그레이드됨
- MXNet이 1.5.1로 업그레이드됨
- Phoenix가 버전 4.14.3으로 업그레이드됨
- Presto가 버전 0.227로 업그레이드됨
- Zeppelin이 버전 0.8.2로 업그레이드됨

새로운 특성

- 이제 Apache Hudi는 클러스터를 생성할 때 Amazon EMR에서 설치할 수 있습니다. 자세한 내용은 Hudi 단원을 참조하십시오.
- (2019년 11월 25일) 이제 클러스터 사용률을 개선하고 비용을 절감하기 위해 여러 단계를 병렬로 실행하도록 선택할 수 있습니다. 대기 중인 단계와 실행 중인 단계를 모두 취소할 수도 있습니다. 자세한 내용은 AWS CLI 및 콘솔을 사용하여 단계 작업을 참조하세요.
- (2019년 12월 3일) 이제 온프레미스 시설에서 네이티브 AWS 서비스, 인프라 및 운영 모델을 AWS Outposts AWS Outposts 활성화하여 EMR 클러스터를 생성하고 실행할 수 있습니다. AWS Outposts 환경에서는 AWS 클라우드에서 사용하는 것과 동일한 AWS APIs, 도구 및 인프라를 사용할 수 있습니다. 자세한 내용은 의 EMR 클러스터를 참조하세요 AWS Outposts.

- (2020년 3월 11일) Amazon EMR 버전 5.28.0부터 AWS 로컬 영역을 지원하는 AWS 리전의 논리적 확장으로 로컬 영역 서브넷에서 Amazon EMR 클러스터를 생성하고 실행할 수 있습니다. 로컬 영역을 사용하면 Amazon EMR 기능과 컴퓨팅 및 스토리지 AWS 서비스와 같은 서비스 하위 집합을 사용자에게 더 가깝게 배치하여 로컬에서 실행되는 애플리케이션에 대한 매우 짧은 지연 시간 액세스를 제공할 수 있습니다. 사용 가능한 로컬 영역 목록은 AWS Local Zones를 참조하세요. 사용 가능한 AWS 로컬 영역에 액세스하는 방법에 대한 자세한 내용은 리전, 가용 영역 및 로컬 영역을 참조하세요.

  로컬 영역은 현재 Amazon EMR Notebooks를 지원하지 않으며 인터페이스 VPC 엔드포인트(AWS PrivateLink)를 사용하여 Amazon EMR에 직접 연결하는 기능도 지원하지 않습니다.

변경 사항, 향상된 기능 및 해결된 문제

- 고가용성 클러스터에 대해 확장된 애플리케이션 지원
  - 자세한 내용은 Amazon EMR 관리 안내서에서 여러 프라이머리 노드가 있는 EMR 클러스터에서 지원되는 애플리케이션을 참조하십시오.
- Spark
  - 성능 최적화
- Hive
  - 성능 최적화
- Presto
  - 성능 최적화

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.

- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.28.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.6	Amazon SageMaker Spark SDK
emr-ddb	4.12.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.11.0	편리한 하둡 에코시스템용 추가 라이브러리

구성 요소	버전	설명
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.13.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.37.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.9.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-hdfs-journalnode	2.8.5-amzn-5	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.10	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.10	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.10	HBase 명령줄 클라이언트
hbase-rest-server	1.4.10	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hbase-thrift-server	1.4.10	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.6-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.6-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.6-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.6-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.6-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.6-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.6-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hudi	0.5.0-incubating	짧은 지연 시간 및 높은 효율성으로 데이터 파이프라인을 구동하는 증분 처리 프레임워크.
hudi-presto	0.5.0-incubating	Hudi로 Presto를 실행하기 위한 번들 라이브러리.
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
jupyterhub	1.0.0	Jupyter 노트북용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.5.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.3-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.3-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.227	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
presto-worker	0.227	여러 쿼리를 실행하는 서비스.
presto-client	0.227	Presto 서버가 시작되지 않은 HA 클러스터의 대기 마스터에 설치된 Presto 명령줄 클라이언트.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.4	Spark 명령줄 클라이언트.
spark-history-server	2.4.4	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.4	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.4	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.2	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.28.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.28.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.

분류	설명
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.27.1

5.27.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, JupyterHubHue, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
AWS Java용 SDK	1.11.615	1.11.615	1.11.595	1.11.566
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.1	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.10	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.5
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.5
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	1.0.0	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.2	4.14.1

	emr-5.27.1	emr-5.27.0	emr-5.26.0	emr-5.25.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.224	0.220	0.220
Spark	2.4.4	2.4.4	2.4.3	2.4.3
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.14.0	1.13.1	1.13.1
Tez	0.9.2	0.9.2	0.9.2	0.9.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.14

5.27.1 릴리스 정보

패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

인스턴스 메타데이터 서비스(IMDS) V2 지원 상태: Amazon EMR 5.23.1, 5.27.1 및 5.32 이상의 구성 요소는 모든 IMDS 직접 호출에 IMDSv2를 사용합니다. 애플리케이션 코드에서 IMDS를 직접 호출하는 경우, IMDSv1 및 IMDSv2를 모두 사용하거나, 추가 보안을 위해 IMDSv2만 사용하도록 IMDS를 구성할 수 있습니다. 다른 5.x EMR 릴리스의 경우 ImDSv1을 비활성화하면 클러스터 시작 오류가 발생합니다.

5.27.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.4	Amazon SageMaker Spark SDK
emr-ddb	4.12.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.11.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.13.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.36.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.8.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.10	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.10	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.10	HBase 명령줄 클라이언트
hbase-rest-server	1.4.10	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.10	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.5-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.5-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.5-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.5-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.5-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.5-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.3.5-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.0.0	Jupyter 노트북용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.4.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.2-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.2-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.224	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.224	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.4	Spark 명령줄 클라이언트.
spark-history-server	2.4.4	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.4	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.4	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.27.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.27.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.

분류	설명
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.

분류	설명
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.

분류	설명
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.27.0

5.27.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, MXNet, Livy, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
AWS Java용 SDK	1.11.615	1.11.595	1.11.566	1.11.546
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.1	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.10	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.5	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.5	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	1.0.0	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.2	4.14.1	4.14.1

	emr-5.27.0	emr-5.26.0	emr-5.25.0	emr-5.24.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.224	0.220	0.220	0.219
Spark	2.4.4	2.4.3	2.4.3	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.14.0	1.13.1	1.13.1	1.12.0
Tez	0.9.2	0.9.2	0.9.2	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.14	3.4.13

5.27.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.27.0에 대한 정보가 포함됩니다. 변경 사항은 5.26.0에 관련됩니다.

최초 릴리스 날짜: 2019년 9월 23일

업그레이드

- AWS SDK for Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- TensorFlow 1.14.0
- 커넥터 및 드라이버:
  - DynamoDB 커넥터 4.12.0

새로운 특성

- (2019년 10월 24일) 다음과 같은 EMR 노트북의 새로운 기능은 모든 Amazon EMR 릴리스에서 사용할 수 있습니다.
  - 이제 Git 리포지토리를 EMR 노트북과 연결하여 버전 제어 환경에서 노트북을 저장할 수 있습니다. 원격 Git 리포지토리를 통해 동료와 코드를 공유할 수 있고 기존 Jupyter Notebook을 재사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon EMR Notebooks에 Git 리포지토리 연결을 참조하세요.
  - 이제 EMR 노트북에서 nbdime 유틸리티를 사용하여 노트북 비교 및 병합을 간소화할 수 있습니다.
  - 이제 EMR 노트북이 JupyterLab을 지원합니다. JupyterLab은 Jupyter Notebook과 완전히 호환 가능한 웹 기반 대화식 개발 환경입니다. 이제 JupyterLab 또는 Jupyter Notebook 편집기에서 노트북을 열 수 있는 옵션이 제공됩니다.
- (2019년 10월 30일) Amazon EMR 버전 5.25.0 이상부터 콘솔에서 클러스터 요약 페이지 또는 애플리케이션 기록 탭을 사용해 Spark 기록 서버 UI에 연결할 수 있습니다. SSH 연결을 통해 웹 프록시를 설정하는 대신, Spark 기록 서버 UI에 액세스하여 애플리케이션 지표를 보고 활성 클러스터 및 종료된 클러스터에 대한 관련 로그 파일에 액세스할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 영구 애플리케이션 사용자 인터페이스에 대해 클러스터 외부에서 액세스를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터
  - 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에 Flink를 설치하고 실행할 수 있습니다. 자세한 내용은 지원되는 애플리케이션 및 기능을 참조하세요.
  - 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에서 HDFS 투명 암호화를 구성할 수 있습니다. 자세한 내용은 여러 프라이머리 노드를 포함하는 EMR 클러스터에서 HDFS의 투명한 암호화를 참조하세요.
  - 이제 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에서 실행 중인 애플리케이션의 구성을 수정할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.
- Amazon EMR-DynamoDB 커넥터
  - 이제 Amazon EMR-DynamoDB 커넥터에서 부울, 목록, 맵, 항목, null과 같은 DynamoDB 데이터 유형을 지원합니다. 자세한 정보는 Hive 테이블을 설정하여 Hive 명령 실행을 참조하십시오.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.27.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.4	Amazon SageMaker Spark SDK
emr-ddb	4.12.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.11.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.13.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.36.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.8.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.8.5-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	1.4.10	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.10	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.10	HBase 명령줄 클라이언트
hbase-rest-server	1.4.10	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.10	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.5-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.5-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.5-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.5-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.5-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.5-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.5-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스

구성 요소	버전	설명
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	1.0.0	Jupyter 노트북용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.4.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.2-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.2-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.224	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.224	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.4	Spark 명령줄 클라이언트.
spark-history-server	2.4.4	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.4	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.4	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.14.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.27.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.27.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.

분류	설명
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.26.0

5.26.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, MXNet, Livy, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
AWS Java용 SDK	1.11.595	1.11.566	1.11.546	1.11.546
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonClo udWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.8.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.10	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.5	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.5	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.4.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.6
Livy	0.6.0	0.6.0	0.6.0	0.6.0
MXNet	1.4.0	1.4.0	1.4.0	1.4.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.2	4.14.1	4.14.1	4.14.1

	emr-5.26.0	emr-5.25.0	emr-5.24.1	emr-5.24.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.220	0.219	0.219
Spark	2.4.3	2.4.3	2.4.2	2.4.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.13.1	1.12.0	1.12.0
Tez	0.9.2	0.9.2	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.14	3.4.13	3.4.13

5.26.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.26.0에 대한 정보가 포함됩니다. 변경 사항은 5.25.0에 관련됩니다.

최초 릴리스 날짜: 2019년 8월 8일

최종 업데이트 날짜: 2019년 8월 19일

업그레이드

- AWS SDK for Java 1.11.595
- HBase 1.4.10
- Phoenix 4.14.2
- 커넥터 및 드라이버:
  - DynamoDB 커넥터 4.11.0
  - MariaDB 커넥터 2.4.2
  - Amazon Redshift JDBC 드라이버 1.2.32.1056

새로운 특성

- (베타) Amazon EMR 5.26.0에서 Lake Formation과 통합된 클러스터를 시작할 수 있습니다. 이 통합은 AWS Glue 데이터 카탈로그의 데이터베이스 및 테이블에 대한 세분화된 열 수준 액세스를 제공합니다. 또한 엔터프라이즈 자격 증명 시스템에서 EMR 노트북 또는 Apache Zeppelin에 연동된 Single Sign-On을 사용할 수도 있습니다. 자세한 내용은 Integrating Amazon EMR with AWS Lake Formation (Beta)를 참조하세요.
- (2019년 8월 19일) 보안 그룹을 지원하는 모든 Amazon EMR 릴리스에 Amazon EMR 퍼블릭 액세스 차단 기능을 사용할 수 있습니다. 퍼블릭 액세스 차단은 각 AWS 리전에 적용되는 계정 전체의 설정입니다. 포트가 예외로 지정되지 않은 한, 퍼블릭 액세스 차단은 클러스터와 연결된 보안 그룹에 포트의 IPv4 0.0.0.0/0 또는 IPv6 ::/0(퍼블릭 액세스)에서 인바운드 트래픽을 허용하는 규칙이 있을 때 클러스터가 시작되지 않도록 합니다. 포트 22는 기본적으로 예외로 설정됩니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon EMR 퍼블랙 액세스 차단 기능 사용을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- EMR Notebooks
  - EMR 5.26.0 이상에서 EMR Notebooks는 기본 Python 라이브러리 외에도 노트북 범위의 Python 라이브러리를 지원합니다. 클러스터를 다시 생성하거나 노트북을 클러스터에 다시 연결할 필요 없이 노트북 편집기 내에서 노트북 범위 라이브러리를 설치할 수 있습니다. 노트북 범위의 라이브러리는 Python 가상 환경에서 생성되므로 현재 노트북 세션에만 적용됩니다. 이를 통해 노트북 종속성을 분리할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 노트북 범위 라이브러리 사용을 참조하세요.
- EMRFS
  - `fs.s3.consistent.metadata.etag.verification.enabled`를 `true`로 설정하여 ETag 확인 기능(베타)을 활성화할 수 있습니다. 이 기능을 통해 EMRFS는 Amazon S3 ETag를 사용하여 읽고 있는 객체가 사용 가능한 최신 버전인지 확인합니다. 이 기능은 동일한 이름을 유지하면서 Amazon S3의 파일을 덮어쓰는 업데이트 후 읽기 사용 사례에 유용합니다. 이 ETag 확인 기능은 현재 S3 Select에서 작동하지 않습니다. 자세한 내용은 일관된 보기 구성 단원을 참조하십시오.
- Spark
  - 동적 파티션 잘라내기, INTERSECT에 앞선 DISTINCT, JOIN-DISINCT 쿼리에 대한 SQL 계획 통계 추론의 개선, 스칼라 하위 쿼리 평면화, 최적화된 조인 재정렬 및 블룸 필터 조인과 같은 최적화가 이제 기본적으로 설정됩니다. 자세한 내용은 Spark 성능 최적화를 참조하십시오.
  - 정렬 병합 조인의 전체 스테이지 코드 생성 기능이 개선되었습니다.
  - 쿼리 조각 및 하위 쿼리 재사용이 개선되었습니다.

  - Spark 시작 시 실행기를 사전 할당하도록 개선되었습니다.
  - 조인의 작은 쪽에 브로드캐스트 힌트가 포함되어 있으면 블룸 필터 조인이 더 이상 적용되지 않습니다.
- Tez의 Hive 실행 시간 비교
  - Tez 관련 문제를 해결했습니다. 이제 Tez UI는 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에서 작동합니다.

알려진 문제

- 정렬 병합 조인에 대한 향상된 전체 스테이지 코드 생성 기능을 사용하면 메모리 압력을 높일 수 있습니다. 이 최적화는 성능을 향상시키지만 `spark.yarn.executor.memoryOverheadFactor`가 충분한 메모리를 제공하도록 조정되지 않은 경우 작업 재시도 또는 실패를 초래할 수 있습니다. 이 기능을 비활성화하려면 `spark.sql.sortMergeJoinExec.extendedCodegen.enabled`를 false로 설정하십시오.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

  ```
  kinit -kt <keytab_file> <principal>
  ```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.26.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.4	Amazon SageMaker Spark SDK
emr-ddb	4.11.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.10.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.12.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.35.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.8.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	2.8.5-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.10	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.10	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.10	HBase 명령줄 클라이언트
hbase-rest-server	1.4.10	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.10	HBase에 Thrift 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hcatalog-client	2.3.5-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.5-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.5-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.5-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.5-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.5-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.5-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.6	Jupyter Notebook용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.

구성 요소	버전	설명
mxnet	1.4.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.2-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.2-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.220	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.220	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.3	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.4.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.13.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.26.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.26.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.

분류	설명
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.25.0

5.25.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Mahout, LivyMXNet, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
AWS Java용 SDK	1.11.566	1.11.546	1.11.546	1.11.519
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.8.0	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.5	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.5	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.4.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.6	0.9.4
Livy	0.6.0	0.6.0	0.6.0	0.5.0
MXNet	1.4.0	1.4.0	1.4.0	1.3.1

	emr-5.25.0	emr-5.24.1	emr-5.24.0	emr-5.23.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.220	0.219	0.219	0.215
Spark	2.4.3	2.4.2	2.4.2	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.13.1	1.12.0	1.12.0	1.12.0
Tez	0.9.2	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.14	3.4.13	3.4.13	3.4.13

5.25.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.25.0에 대한 정보가 포함됩니다. 변경 사항은 5.24.1에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 17일

최종 업데이트 날짜: 2019년 10월 30일

Amazon EMR 5.25.0

업그레이드

- AWS SDK for Java 1.11.566

- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

새로운 특성

- (2019년 10월 30일) Amazon EMR 버전 5.25.0부터 콘솔에서 클러스터 요약 페이지 또는 애플리케이션 기록 탭을 사용해 Spark 기록 서버 UI에 연결할 수 있습니다. SSH 연결을 통해 웹 프록시를 설정하는 대신, Spark 기록 서버 UI에 액세스하여 애플리케이션 지표를 보고 활성 클러스터 및 종료된 클러스터에 대한 관련 로그 파일에 액세스할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 영구 애플리케이션 사용자 인터페이스에 대해 클러스터 외부에서 액세스를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Bloom 필터로 입력을 사전 필터링하여 일부 조인의 성능을 개선했습니다. 이 최적화는 기본적으로 비활성화되어 있으며, Spark 구성 파라미터 `spark.sql.bloomFilterJoin.enabled`를 `true`로 설정하여 활성화할 수 있습니다.
  - 문자열 형식 열을 기준으로 그룹화하는 성능을 개선했습니다.
  - HBase를 설치하지 않고 클러스터에 대한 R4 인스턴스 유형의 핵심 구성 및 기본 Spark 실행기 메모리를 개선했습니다.
  - 잘린 테이블이 조인의 왼쪽에 있어야 하는 이전의 동적 파티션 잘라내기 기능에 대한 문제를 해결했습니다.
  - 별칭과 관련된 추가 사례를 적용하기 위해 INTERSECT 이전 DISTINCT 최적화를 개선했습니다.
  - DISTINCT 쿼리 이전 JOIN에 대한 SQL 계획 통계 추론을 개선했습니다. 이 개선 사항은 기본적으로 비활성화되어 있으며, Spark 구성 파라미터 `spark.sql.statsImprovements.enabled`를 `true`로 설정하여 활성화할 수 있습니다. 이 최적화는 Intersect 이전 Distinct 기능에 필요하며 `spark.sql.optimizer.distinctBeforeIntersect.enabled`를 `true`로 설정할 경우 자동으로 활성화됩니다.

  - 테이블 크기 및 필터에 따라 조인 순서를 최적화했습니다. 이 최적화는 기본적으로 비활성화되어 있으며, Spark 구성 파라미터 `spark.sql.optimizer.sizeBasedJoinReorder.enabled`를 `true`로 설정하여 활성화할 수 있습니다.

  자세한 내용은 Spark 성능 최적화를 참조하십시오.

- EMRFS
  - 이제 EMRFS 설정 `fs.s3.buckets.create.enabled`가 기본적으로 비활성화됩니다. 테스트 결과 이 설정을 비활성화하면 성능이 향상되고 의도치 않은 S3 버킷 생성이 방지된다는 사실을 확인했습니다. 애플리케이션이 이 기능을 사용하는 경우 `emrfs-site` 구성 분류에서 속성 `fs.s3.buckets.create.enabled`를 `true`로 설정하여 활성화할 수 있습니다. 자세한 내용은 클러스터를 생성할 때 구성 제공 단원을 참조하십시오.
- 보안 구성의 로컬 디스크 암호화 및 S3 암호화 개선(2019년 8월 5일)
  - 보안 구성 설정에서 로컬 디스크 암호화 설정과 Amazon S3 암호화 설정이 분리되었습니다.
  - EBS 암호화를 활성화하는 옵션이 5.24.0 이상 릴리스에 추가되었습니다. 이 옵션을 선택하면 스토리지 볼륨뿐만 아니라 루트 디바이스 볼륨도 암호화됩니다. 이전 버전에서는 사용자 지정 AMI를 사용하여 루트 디바이스 볼륨을 암호화해야 했습니다.
  - 자세한 내용은 Amazon EMR 관리 안내서에서 암호화 옵션을 참조하세요.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

  ```
  kinit -kt <keytab_file> <principal>
  ```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.25.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.4	Amazon SageMaker Spark SDK
emr-ddb	4.10.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.9.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.34.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.8.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.9	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.9	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.9	HBase 명령줄 클라이언트
hbase-rest-server	1.4.9	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hbase-thrift-server	1.4.9	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.5-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.5-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.5-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.5-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.5-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.5-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.5-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.6	Jupyter Notebook용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.4.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.1-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.1-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.220	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.220	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.3	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.4.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.13.1	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.2	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.14	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.14	ZooKeeper 명령줄 클라이언트.

5.25.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.25.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
recordserver-env	EMR RecordServer 환경의 값을 변경합니다.
recordserver-conf	EMR RecordServer의 erver.properties 파일의 값을 변경합니다.
recordserver-log4j	EMR RecordServer의 log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.

분류	설명
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.24.1

5.24.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, Mahout, MXNetOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
AWS Java용 SDK	1.11.546	1.11.546	1.11.519	1.11.519
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.8.0	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.4.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.6	0.9.4	0.9.4
Livy	0.6.0	0.6.0	0.5.0	0.5.0
MXNet	1.4.0	1.4.0	1.3.1	1.3.1

	emr-5.24.1	emr-5.24.0	emr-5.23.1	emr-5.23.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.219	0.215	0.215
Spark	2.4.2	2.4.2	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.24.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.24.1에 대한 정보가 포함됩니다. 변경 사항은 5.24.0에 관련됩니다.

최초 릴리스 날짜: 2019년 6월 26일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

  > **Note**
  >
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.24.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.9.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.8.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.33.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.8.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.8.5-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	1.4.9	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.9	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.9	HBase 명령줄 클라이언트
hbase-rest-server	1.4.9	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.9	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스

구성 요소	버전	설명
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.6	Jupyter Notebook용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.4.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.1-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.1-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.219	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.219	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.2	Spark 명령줄 클라이언트.
spark-history-server	2.4.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.24.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.24.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.

분류	설명
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.24.0

5.24.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNetMahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
AWS Java용 SDK	1.11.546	1.11.519	1.11.519	1.11.510
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.8.0	1.7.1	1.7.1	1.7.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.9
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.4.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-

	emr-5.24.0	emr-5.23.1	emr-5.23.0	emr-5.22.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.6	0.9.4	0.9.4	0.9.4
Livy	0.6.0	0.5.0	0.5.0	0.5.0
MXNet	1.4.0	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.1.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.1
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.219	0.215	0.215	0.215
Spark	2.4.2	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.1
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.24.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.24.0에 대한 정보가 포함됩니다. 변경 사항은 5.23.0에 관련됩니다.

최초 릴리스 날짜: 2019년 6월 11일

최종 업데이트 날짜: 2019년 8월 5일

업그레이드

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0.9.6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK for Java 1.11.546
- 커넥터 및 드라이버:
  - DynamoDB 커넥터 4.9.0
  - MariaDB 커넥터 2.4.1
  - Amazon Redshift JDBC 드라이버 1.2.27.1051

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - 파티션을 동적으로 삭제하는 최적화가 추가되었습니다. 최적화는 기본적으로 비활성화되어 있습니다. 활성화하려면 Spark 구성 파라미터 `spark.sql.dynamicPartitionPruning.enabled`를 `true`로 설정하십시오.
  - `INTERSECT` 쿼리의 성능을 개선했습니다. 이 최적화는 기본적으로 비활성화되어 있습니다. 활성화하려면 Spark 구성 파라미터 `spark.sql.optimizer.distinctBeforeIntersect.enabled`를 `true`로 설정하십시오.
  - 동일한 관계를 사용하는 집계로 스칼라 하위 쿼리를 평면화하는 최적화가 추가되었습니다. 최적화는 기본적으로 비활성화되어 있습니다. 활성화하려면 Spark 구성 파라미터 `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled`를 `true`로 설정하십시오.
  - 전체 스테이지 코드 생성이 개선되었습니다.

  자세한 내용은 Spark 성능 최적화를 참조하십시오.

- 보안 구성의 로컬 디스크 암호화 및 S3 암호화 개선(2019년 8월 5일)
  - 보안 구성 설정에서 로컬 디스크 암호화 설정과 Amazon S3 암호화 설정이 분리되었습니다.
  - EBS 암호화를 활성화하는 옵션이 추가되었습니다. 이 옵션을 선택하면 스토리지 볼륨뿐만 아니라 루트 디바이스 볼륨도 암호화됩니다. 이전 버전에서는 사용자 지정 AMI를 사용하여 루트 디바이스 볼륨을 암호화해야 했습니다.
  - 자세한 내용은 Amazon EMR 관리 안내서에서 [암호화 옵션](#)을 참조하세요.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.24.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.9.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.8.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.3.0	EMR S3Select 커넥터
emrfs	2.33.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.8.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-4	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.8.5-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.9	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.9	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.9	HBase 명령줄 클라이언트
hbase-rest-server	1.4.9	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.9	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.4-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.4.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.6	Jupyter Notebook용 다중 사용자 서버
livy-server	0.6.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.4.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구키트

구성 요소	버전	설명
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.1-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.1-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.219	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.219	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.2	Spark 명령줄 클라이언트.
spark-history-server	2.4.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.

구성 요소	버전	설명
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.24.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.24.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.

분류	설명
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.

분류	설명
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.

분류	설명
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.23.1

5.23.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hue, Hive, JupyterHub, Livy, Mahout, MXNetOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
AWS Java용 SDK	1.11.519	1.11.519	1.11.510	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.1	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.9	1.4.8

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.1.0	5.0.0
Phoenix	4.14.1	4.14.1	4.14.1	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1

	emr-5.23.1	emr-5.23.0	emr-5.22.0	emr-5.21.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.1	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.23.1 릴리스 정보

패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

인스턴스 메타데이터 서비스(IMDS) V2 지원 상태: Amazon EMR 5.23.1, 5.27.1 및 5.32 이상의 구성 요소는 모든 IMDS 직접 호출에 IMDSv2를 사용합니다. 애플리케이션 코드에서 IMDS를 직접 호출하는 경우, IMDSv1 및 IMDSv2를 모두 사용하거나, 추가 보안을 위해 IMDSv2만 사용하도록 IMDS를 구성할 수 있습니다. 다른 5.x EMR 릴리스의 경우 ImDSv1을 비활성화하면 클러스터 시작 오류가 발생합니다.

5.23.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.8.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.7.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.32.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.7.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.8.5-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-3	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.9	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.9	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	1.4.9	HBase 명령줄 클라이언트
hbase-rest-server	1.4.9	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.9	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버

구성 요소	버전	설명
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.1-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.1-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.215	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.215	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.0	Spark 명령줄 클라이언트.
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.23.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.23.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.

분류	설명
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.

분류	설명
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.23.0

5.23.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
AWS Java용 SDK	1.11.519	1.11.510	1.11.479	1.11.479

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.1	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.9	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.1.0	5.0.0	5.0.0

	emr-5.23.0	emr-5.22.0	emr-5.21.2	emr-5.21.1
Phoenix	4.14.1	4.14.1	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.1	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.23.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.23.0에 대한 정보가 포함됩니다. 변경 사항은 5.22.0에 관련됩니다.

최초 릴리스 날짜: 2019년 4월 1일

최종 업데이트 날짜: 2019년 4월 30일

업그레이드

- AWS SDK for Java 1.11.519

새로운 특성

- (2019년 4월 30일) Amazon EMR 5.23.0 이상에서 3개의 프라이머리 노드를 포함하는 클러스터를 시작하고 YARN Resource Manager, HDFS NameNode, Spark, Hive, Ganglia와 같은 애플리케이션의 고가용성을 지원할 수 있습니다. 이 기능을 사용하면 프라이머리 노드가 더 이상 잠재적 단일 장

애 지점이 아닙니다. 프라이머리 노드 중 하나에 장애가 발생할 경우, Amazon EMR이 자동으로 대기 프라이머리 노드로 장애 조치하고 장애가 발생한 프라이머리 노드를 동일한 구성 및 부트스트랩 작업을 갖는 새로운 프라이머리 노드로 교체합니다. 자세한 내용은 프라이머리 노드 계획 및 구성을 참조하세요.

알려진 문제

- Tez UI(Amazon EMR 릴리스 5.26.0에서 수정됨)

  Tez UI는 여러 프라이머리 노드를 포함하는 EMR 클러스터에서 작동하지 않습니다.
- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.
    2. 텍스트 편집기를 사용하여 hue.ini 파일을 엽니다. 예시:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. appblacklist라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

       ```
       appblacklist = search
       ```
    4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

       ```
       sudo stop hue; sudo start hue
       ```
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작

업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.23.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.8.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.7.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분산 복사 애플리케이션.
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.32.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.7.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.8.5-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-hdfs-journalnode	2.8.5-amzn-3	HA 클러스터에서 하둡 파일 시스템 저널을 관리하기 위한 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.9	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.4.9	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.9	HBase 명령줄 클라이언트
hbase-rest-server	1.4.9	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.9	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.1-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.1-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.215	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
presto-worker	0.215	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.0	Spark 명령줄 클라이언트.
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.23.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.23.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.

분류	설명
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.

분류	설명
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.

분류	설명
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.

분류	설명
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.22.0

5.22.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HueJupyterHub, MXNet, Livy, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
AWS Java용 SDK	1.11.510	1.11.479	1.11.479	1.11.479
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.1	1.7.0	1.7.0	1.7.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.9	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.22.0	emr-5.21.2	emr-5.21.1	emr-5.21.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.1.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.1	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.215
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.1	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.22.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.22.0에 대한 정보가 포함됩니다. 변경 사항은 5.21.0에 관련됩니다.

> **Important**
>
> Amazon EMR 릴리스 5.22.0부터 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 이전 Amazon EMR 릴리스는 경우에 따라 AWS 서명 버전 4만 사용된다는 릴리스 정보가 표시되지 않는 한 서명 버전 2를 사용합니다. 자세한 내용은 Amazon Simple Storage Service 개발자 안내서의 요청 인증(AWS 서명 버전 4) 및 요청 인증(AWS 서명 버전 2)을 참조하세요.

최초 릴리스 날짜: 2019년 3월 20일

업그레이드

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- 커넥터 및 드라이버:
  - DynamoDB 커넥터 4.8.0
  - MariaDB Connector 2.2.6
  - Amazon Redshift JDBC 드라이버 1.2.20.1043

새로운 특성

- EBS 전용 스토리지를 사용하는 EC2 인스턴스 유형에 대한 기본 EBS 구성을 수정했습니다. Amazon EMR 릴리스 5.22.0 이상을 사용하여 클러스터를 생성할 때 EBS 스토리지의 기본적인 크기가 인스턴스 크기에 따라 증가합니다. 뿐만 아니라 증가한 스토리지를 여러 볼륨에 분할하여 향상된 IOPS 성능을 제공합니다. 다른 EBS 인스턴스 스토리지 구성을 사용하려는 경우 EMR 클러스터를 생성하거나 기존 클러스터에 노드를 추가할 때 이 구성을 지정할 수 있습니다. 각 인스턴스 유형에 기본적으로 할당된 스토리지 크기와 볼륨 수에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 인스턴스의 기본 EBS 스토리지를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - YARN의 Spark에 새로운 구성 속성 `spark.yarn.executor.memoryOverheadFactor`를 추가했습니다. 이 속성의 값은 최소 384MB의 메모리 오버헤드 값을 실행기 메모리의 백분율로 설정하는 확장 팩터입니다. 메모리 오버헤드가 `spark.yarn.executor.memoryOverhead`를 사용하여 명시적으로 설정된 경우 이 속성은 아무 효과가 없습니다. 기본값은 18.75%를 나타내는 `0.1875`입니다. Amazon EMR의 기본값은 실행기 메모리 오버헤드를 위해 Spark에서 내부적으로 설정한 기본값 10%보다 더 많은 공간을 YARN 컨테이너에 남깁니다. TPC-DS 벤치마크 결과 Amazon EMR 기본값, 18.75%로 설정하면 메모리 관련 오류가 경험적으로 적었습니다.
  - 성능 개선을 위해 SPARK-26316을 백포트했습니다.

- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수 있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.

알려진 문제

- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.
    2. 텍스트 편집기를 사용하여 `hue.ini` 파일을 엽니다. 예시:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. `appblacklist`라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

       ```
       appblacklist = search
       ```

    4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

       ```
       sudo stop hue; sudo start hue
       ```

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.22.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.8.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.6.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.31.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.7.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.8.5-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.9	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.9	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.9	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.4.9	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.9	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스

구성 요소	버전	설명
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.1.0	Oozie 명령줄 클라이언트
oozie-server	5.1.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.1-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.1-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.215	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.215	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트

구성 요소	버전	설명
spark-client	2.4.0	Spark 명령줄 클라이언트.
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.22.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.22.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.

분류	설명
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.

분류	설명
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.21.2

5.21.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, JupyterHubHue, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
AWS Java용 SDK	1.11.479	1.11.479	1.11.479	1.11.461

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonClo udWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.7.0	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0

	emr-5.21.2	emr-5.21.1	emr-5.21.0	emr-5.20.1
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.215	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.21.2 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.21.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.1	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.30.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.7.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.8	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.8	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.8	HBase 명령줄 클라이언트
hbase-rest-server	1.4.8	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.8	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.3.4-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.215	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.215	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.0	Spark 명령줄 클라이언트.
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.21.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.21.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.

분류	설명
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.

분류	설명
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.

분류	설명
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.21.1

5.21.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
AWS Java용 SDK	1.11.479	1.11.479	1.11.461	1.11.461
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.6
Scala	2.11.12	2.11.12	2.11.12	2.11.12
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.7.0	1.6.2	1.6.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.8
HCatalog	2.3.4	2.3.4	2.3.4	2.3.4
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.4
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.3.0
Iceberg	-	-	-	-

	emr-5.21.1	emr-5.21.0	emr-5.20.1	emr-5.20.0
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.215	0.214	0.214
Spark	2.4.0	2.4.0	2.4.0	2.4.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.12.0
Tez	0.9.1	0.9.1	0.9.1	0.9.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.21.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.21.1에 대한 정보가 포함됩니다. 변경 사항은 5.21.0에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 18일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

  > **Note**
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.21.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을

위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.1	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.30.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.7.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.8	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.8	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.8	HBase 명령줄 클라이언트
hbase-rest-server	1.4.8	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.8	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.3.4-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.215	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.215	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.0	Spark 명령줄 클라이언트.
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.21.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.21.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.

분류	설명
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.

분류	설명
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.

분류	설명
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.21.0

5.21.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
AWS Java용 SDK	1.11.479	1.11.461	1.11.461	1.11.433
Python	2.7, 3.6	2.7, 3.6	2.7, 3.6	2.7, 3.4
Scala	2.11.12	2.11.12	2.11.12	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.7.0	1.6.2	1.6.2	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.8	1.4.7
HCatalog	2.3.4	2.3.4	2.3.4	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.4	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.3.0	4.2.0
Iceberg	-	-	-	-

	emr-5.21.0	emr-5.20.1	emr-5.20.0	emr-5.19.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.1	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.215	0.214	0.214	0.212
Spark	2.4.0	2.4.0	2.4.0	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.12.0	1.11.0
Tez	0.9.1	0.9.1	0.9.1	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.21.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.21.0에 대한 정보가 포함됩니다. 변경 사항은 5.20.0에 관련됩니다.

최초 릴리스 날짜: 2019년 2월 18일

최종 업데이트 날짜: 2019년 4월 3일

업그레이드

- Flink 1.7.0
- Presto 0.215
- AWS SDK for Java 1.11.479

새로운 특성

- (2019년 4월 3일) Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- Zeppelin
  - ZEPPELIN-3878을 백포트했습니다.

알려진 문제

- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.

2. 텍스트 편집기를 사용하여 hue.ini 파일을 엽니다. 예시:

```
sudo vim /etc/hue/conf/hue.ini
```

3. appblacklist라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

```
appblacklist = search
```

4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

```
sudo stop hue; sudo start hue
```

- Tez의 Hive 실행 시간 비교
  - 이 문제는 Amazon EMR 5.22.0에서 해결되었습니다.

    클러스터 프라이머리 노드에 대한 SSH 연결을 통해 http://*MasterDNS*:8080/tez-ui의 Tez UI에 연결할 때 'Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled(어댑터 작동이 실패했습니다. - 타임라인 서버(ATS)가 범위를 벗어났습니다. 다운되었거나 CORS가 활성화되어 있지 않습니다.)"라는 오류가 나타나거나, 작업 중 예기치 않게 N/A(해당 사항 없음)가 표시됩니다.

    이 오류는 Tez UI가 프라이머리 노드의 호스트 이름 대신 `localhost`를 사용하여 YARN 타임라인 서버에 요청할 때 발생합니다. 이 문제를 해결하기 위해 스크립트를 부트스트랩 작업 또는 단계로 실행할 수 있습니다. 스크립트는 Tez `configs.env` 파일의 호스트 이름을 업데이트합니다. 스크립트에 대한 자세한 내용과 위치를 보려면 부트스트랩 지침을 참조하십시오.
- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수 있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.21.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터

구성 요소	버전	설명
emr-goodies	2.5.1	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.11.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.30.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.7.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리

구성 요소	버전	설명
hadoop-hdfs-namenode	2.8.5-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.8	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.8	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.8	HBase 명령줄 클라이언트
hbase-rest-server	1.4.8	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hbase-thrift-server	1.4.8	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.215	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.215	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.21.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.21.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.20.1

5.20.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
AWS Java용 SDK	1.11.461	1.11.461	1.11.433	1.11.433
Python	2.7, 3.6	2.7, 3.6	2.7, 3.4	2.7, 3.4

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
Scala	2.11.12	2.11.12	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.2	1.6.1	1.6.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.8	1.4.7	1.4.7
HCatalog	2.3.4	2.3.4	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.5
Hive	2.3.4	2.3.4	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.3.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.9.4
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.1	1.3.0	1.3.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0

	emr-5.20.1	emr-5.20.0	emr-5.19.1	emr-5.19.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.214	0.212	0.212
Spark	2.4.0	2.4.0	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.12.0	1.11.0	1.11.0
Tez	0.9.1	0.9.1	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.13

5.20.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.20.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.1	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.29.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.6.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	1.4.8	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.8	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.8	HBase 명령줄 클라이언트
hbase-rest-server	1.4.8	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.8	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.4-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스

구성 요소	버전	설명
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.214	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.214	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.0	Spark 명령줄 클라이언트.
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.20.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.20.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.

분류	설명
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.20.0

5.20.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
AWS Java용 SDK	1.11.461	1.11.433	1.11.433	1.11.393
Python	2.7, 3.6	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.12	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.2	1.6.1	1.6.1	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.8	1.4.7	1.4.7	1.4.7
HCatalog	2.3.4	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.5	2.8.4
Hive	2.3.4	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.3.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.20.0	emr-5.19.1	emr-5.19.0	emr-5.18.1
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.9.4	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.1	1.3.0	1.3.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.214	0.212	0.212	0.210
Spark	2.4.0	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.12.0	1.11.0	1.11.0	1.9.0
Tez	0.9.1	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.13	3.4.12

5.20.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.20.0에 대한 정보가 포함됩니다. 변경 사항은 5.19.0에 관련됩니다.

최초 릴리스 날짜: 2018년 12월 18일

마지막 업데이트 날짜: 2019년 1월 22일

업그레이드

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK for Java 1.11.461

새로운 특성

- (2019년 1월 22일) Amazon EMR의 Kerberos가 외부 KDC의 인증 보안 주체를 지원하도록 개선되었습니다. 이는 여러 클러스터가 하나의 외부 KDC를 공유할 수 있기 때문에 보안 주체 관리를 중앙 집중화합니다. 또한 외부 KDC는 Active Directory 도메인과 교차 영역 신뢰를 가질 수 있습니다. 이를 통해 모든 클러스터는 Active Directory의 보안 주체를 인증할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Kerberos 인증 사용을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR의 기본 Amazon Linux AMI
  - Python3 패키지가 python 3.4에서 3.6으로 업그레이드되었습니다.
- EMRFS S3 최적화 커미터
  - EMRFS S3 최적화 커미터가 이제 기본적으로 활성화되며 쓰기 성능을 향상시킵니다. 자세한 내용은 EMRFS S3 최적화 커미터 사용 단원을 참조하십시오.
- Hive
  - HIVE-16686를 백포트했습니다.

- Glue와 Spark 및 Hive
  - EMR 5.20.0 이상에서는 AWS Glue 데이터 카탈로그를 메타스토어로 사용할 때 Spark 및 Hive에 대해 병렬 파티션 정리가 자동으로 활성화됩니다. 이 변경 사항은 여러 요청을 병렬로 실행하여 파티션을 검색함으로써 쿼리 계획 시간을 크게 단축합니다. 동시에 실행할 수 있는 총 세그먼트 수는 1~10입니다. 기본값은 5이며, 또한 권장 설정입니다. `hive-site` 구성 분류에서 `aws.glue.partition.num.segments` 속성을 지정하여 이 설정을 변경할 수 있습니다. 조절이 발생할 경우 값을 1로 변경하여 기능을 끌 수 있습니다. 자세한 내용은 [AWS Glue 그먼트 구조]를 참조하세요.

알려진 문제

- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.
    2. 텍스트 편집기를 사용하여 `hue.ini` 파일을 엽니다. 예시:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. `appblacklist`라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

       ```
       appblacklist = search
       ```

    4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

       ```
       sudo stop hue; sudo start hue
       ```

- Tez의 Hive 실행 시간 비교
  - 이 문제는 Amazon EMR 5.22.0에서 해결되었습니다.

클러스터 프라이머리 노드에 대한 SSH 연결을 통해 http://*MasterDNS*:8080/tez-ui의 Tez UI에 연결할 때 'Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled(어댑터 작동이 실패했습니다. - 타임라인 서버(ATS)가 범위를 벗어났습니다. 다운되었거나 CORS가 활성화되어 있지 않습니다.)"라는 오류가 나타나거나, 작업 중 예기치 않게 N/A(해당 사항 없음)가 표시됩니다.

이 오류는 Tez UI가 프라이머리 노드의 호스트 이름 대신 `localhost`를 사용하여 YARN 타임라인 서버에 요청할 때 발생합니다. 이 문제를 해결하기 위해 스크립트를 부트스트랩 작업 또는 단계로 실행할 수 있습니다. 스크립트는 Tez `configs.env` 파일의 호스트 이름을 업데이트합니다. 스크립트에 대한 자세한 내용과 위치를 보려면 부트스트랩 지침을 참조하십시오.

- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수 있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 `hadoop` 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 `hadoop` 사용자의 Kerberos 티켓을 갱신합니다.

  ```
  kinit -kt <keytab_file> <principal>
  ```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

5.20.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.1	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.1	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emr-s3-select	1.2.0	EMR S3Select 커넥터
emrfs	2.29.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.6.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.8.5-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.8	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.8	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.8	HBase 명령줄 클라이언트
hbase-rest-server	1.4.8	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.8	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.4-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	2.3.4-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.4-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.4-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.4-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.4-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.4-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.3.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.1	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.

구성 요소	버전	설명
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.214	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.214	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.4.0	Spark 명령줄 클라이언트.
spark-history-server	2.4.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.4.0	YARN용 인 메모리 실행 엔진.

구성 요소	버전	설명
spark-yarn-slave	2.4.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.12.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.9.1	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.20.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.20.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.19.1

5.19.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
AWS Java용 SDK	1.11.433	1.11.433	1.11.393	1.11.393
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.1	1.6.0	1.6.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.7
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.5	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.9.4	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.3.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0

	emr-5.19.1	emr-5.19.0	emr-5.18.1	emr-5.18.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.212	0.210	0.210
Spark	2.3.2	2.3.2	2.3.2	2.3.2
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.11.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.8.0
ZooKeeper	3.4.13	3.4.13	3.4.12	3.4.12

5.19.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.19.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.0	Amazon SageMaker Spark SDK
emr-ddb	4.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.1	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.1.0	EMR S3Select 커넥터
emrfs	2.28.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.6.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.5-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	1.4.7	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.7	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.7	HBase 명령줄 클라이언트
hbase-rest-server	1.4.7	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.7	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스

구성 요소	버전	설명
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.212	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.212	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.2	Spark 명령줄 클라이언트.
spark-history-server	2.3.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.19.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.19.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.

분류	설명
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.19.0

5.19.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HueJupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
AWS Java용 SDK	1.11.433	1.11.393	1.11.393	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.1	1.6.0	1.6.0	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.7	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.5	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-

	emr-5.19.0	emr-5.18.1	emr-5.18.0	emr-5.17.2
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.9.4	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.3.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.212	0.210	0.210	0.206
Spark	2.3.2	2.3.2	2.3.2	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.11.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.8.0	0.7.3
ZooKeeper	3.4.13	3.4.12	3.4.12	3.4.12

5.19.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.19.0에 대한 정보가 포함됩니다. 변경 사항은 5.18.0에 관련됩니다.

최초 릴리스 날짜: 2018년 11월 7일

최종 업데이트: 2018년 11월 19일

업그레이드

- 하둡 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- Zookeeper 3.4.13
- AWS SDK for Java 1.11.433

새로운 특성

- (2018년 11월 19일) EMR Notebooks는 Jupyter Notebook 기반의 관리형 환경입니다. PySpark, Spark SQL, Spark R 및 Scala에 대한 Spark 매직 커널을 지원합니다. EMR Notebooks는 Amazon EMR 릴리스 5.18.0 이상을 사용하여 생성한 클러스터와 함께 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [EMR Notebooks 사용]을 참조하세요.
- Spark 및 EMRFS를 사용하여 Parquet 파일을 작성할 때 EMRFS S3 최적화 커미터를 사용할 수 있습니다. 이 커미터는 쓰기 성능을 향상시킵니다. 자세한 내용은 [EMRFS S3 최적화 커미터 사용] 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- YARN
  - 코어 노드에서 실행할 때 애플리케이션 마스터 프로세스를 제한하는 로직이 수정되었습니다. 이 기능은 이제 YARN 노드 레이블 기능과 `yarn-site`의 속성 및 `capacity-scheduler` 구성 분류를 사용합니다. 자세한 정보는 https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN. 섹션을 참조하세요.
- Amazon EMR의 기본 Amazon Linux AMI
  - `ruby18`, `php56`, `gcc48`이 이제 기본적으로 설치되지 않습니다. 필요할 경우 `yum`을 사용하여 설치할 수 있습니다.

- aws-java-sdk ruby gem이 이제 더 이상 기본적으로 설치되지 않습니다. 필요할 경우 `gem install aws-sdk`를 사용하여 설치할 수 있습니다. 특정 구성 요소도 설치할 수 있습니다. 예를 들어 `gem install aws-sdk-s3`입니다.

알려진 문제

- EMR Notebooks - 일부 상황에서 노트북 편집기가 여러 개 열려 있는 경우 노트북 편집기가 클러스터에 연결할 수 없는 것처럼 보일 수 있습니다. 이 경우 브라우저 쿠키를 지우고 노트북 편집기를 다시 여십시오.
- CloudWatch ContainerPending 지표 및 자동 조정 - (5.20.0에서 수정됨) Amazon EMR은 `ContainerPending`에 대해 음수 값을 생성할 수 있습니다. 자동 조정 규칙에 `ContainerPending`을 사용할 경우 자동 조정이 예상대로 작동하지 않습니다. 자동 조정에 `ContainerPending`을 사용하지 마십시오.
- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수 있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.

5.19.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.2.0	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.7.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.1	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.1.0	EMR S3Select 커넥터
emrfs	2.28.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.6.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.5-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.5-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.8.5-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.5-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.5-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.5-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.5-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.5-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.5-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.5-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.7	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.7	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.7	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.4.7	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.7	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.9.4	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스

구성 요소	버전	설명
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.3.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.212	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.212	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트

구성 요소	버전	설명
spark-client	2.3.2	Spark 명령줄 클라이언트.
spark-history-server	2.3.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.11.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.13	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.13	ZooKeeper 명령줄 클라이언트.

5.19.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-5.19.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.

분류	설명
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-memory	Presto memory.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
presto-connector-tpcds	Presto tpcds.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.

분류	설명
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.18.1

5.18.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
AWS Java용 SDK	1.11.393	1.11.393	1.11.336	1.11.336

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.6.0	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.7	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0

	emr-5.18.1	emr-5.18.0	emr-5.17.2	emr-5.17.1
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.210	0.206	0.206
Spark	2.3.2	2.3.2	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.8.0	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

5.18.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.18.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 myapp-component라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 2.2-amzn-2로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.1.0	EMR S3Select 커넥터
emrfs	2.27.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.6.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.4-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.4-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.4-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.4-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.4-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.4-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.4-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.4-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.4-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.7	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.7	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.7	HBase 명령줄 클라이언트
hbase-rest-server	1.4.7	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.7	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.3.3-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.2.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.210	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.210	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.2	Spark 명령줄 클라이언트.
spark-history-server	2.3.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.9.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.18.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.18.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.

분류	설명
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.

분류	설명
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.18.0

5.18.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, LivyMXNet, Mahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
AWS Java용 SDK	1.11.393	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.6.0	1.5.2	1.5.2	1.5.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.7	1.4.6	1.4.6	1.4.6
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1

	emr-5.18.0	emr-5.17.2	emr-5.17.1	emr-5.17.0
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.210	0.206	0.206	0.206
Spark	2.3.2	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	1.9.0
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.8.0	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

5.18.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.18.0에 대한 정보가 포함됩니다. 변경 사항은 5.17.0에 관련됩니다.

최초 릴리스 날짜: 2018년 10월 24일

업그레이드

- Flink 1.6.0

- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

새로운 특성

- Amazon EMR 5.18.0부터 Amazon EMR 아티팩트 리포지토리를 사용하여 특정 Amazon EMR 릴리스 버전에서 사용할 수 있는 정확한 버전의 라이브러리와 종속 항목에 대해 작업 코드를 빌드할 수 있습니다. 자세한 내용은 Amazon EMR 아티팩트 리포지토리를 사용하여 종속성 확인 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- Hive
  - S3 Select에 대한 지원을 추가했습니다. 자세한 내용은 S3 Select와 함께 Hive를 사용하여 성능 향상 단원을 참조하십시오.
- Presto
  - S3 Select Pushdown에 대한 지원을 추가했습니다. 자세한 내용은 S3 Select Pushdown과 함께 Presto를 사용하여 성능 향상 단원을 참조하십시오.
- Spark
  - Spark에 대한 기본 log4j 구성이 Spark 스트리밍 작업에 대한 컨테이너 로그를 1시간마다 넘기도록 변경되었습니다. 따라서 오래 실행되는 Spark 스트리밍 작업에 대한 로그가 삭제되지 않습니다.

5.18.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.1.0	EMR S3Select 커넥터
emrfs	2.27.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.6.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.4-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.4-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.4-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.4-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.4-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.4-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.4-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.4-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.4-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	1.4.7	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.7	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.7	HBase 명령줄 클라이언트
hbase-rest-server	1.4.7	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.7	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스

구성 요소	버전	설명
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
nginx	1.12.1	HTTP인 nginx [engine x] 및 역방향 프록시 서버
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.2.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.210	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.210	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.2	Spark 명령줄 클라이언트.
spark-history-server	2.3.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.9.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.8.0	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.18.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.18.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.17.2

5.17.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HueJupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
AWS Java용 SDK	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.2	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.6	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1

	emr-5.17.2	emr-5.17.1	emr-5.17.0	emr-5.16.1
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.206	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	1.9.0	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

5.17.2 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.17.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.0.0	EMR S3Select 커넥터
emrfs	2.26.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.5.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.4-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.4-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.4-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.4-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.4-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.4-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.4-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.4-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resour cemanager	2.8.4-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.6	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.6	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.6	HBase 명령줄 클라이언트
hbase-rest-server	1.4.6	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.6	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-1	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.3.3-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.2.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.206	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.206	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.1	Spark 명령줄 클라이언트.
spark-history-server	2.3.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.9.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.17.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.17.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.

분류	설명
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.

분류	설명
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.17.1

5.17.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hue, Hive, JupyterHub, MXNet, Livy, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
AWS Java용 SDK	1.11.336	1.11.336	1.11.336	1.11.336
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.2	1.5.0	1.5.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.6	1.4.4	1.4.4
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.4
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1

	emr-5.17.1	emr-5.17.0	emr-5.16.1	emr-5.16.0
Livy	0.5.0	0.5.0	0.5.0	0.5.0
MXNet	1.2.0	1.2.0	1.2.0	1.2.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.14.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.206	0.203	0.203
Spark	2.3.1	2.3.1	2.3.1	2.3.1
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	1.9.0	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

5.17.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.17.1에 대한 정보가 포함됩니다. 변경 사항은 5.17.0에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 18일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

5.17.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK
emr-ddb	4.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.0.0	EMR S3Select 커넥터

구성 요소	버전	설명
emrfs	2.26.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.5.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.4-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.4-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.4-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.4-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.4-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.4-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.8.4-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.4-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.4-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.6	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.6	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.6	HBase 명령줄 클라이언트
hbase-rest-server	1.4.6	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.6	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.2.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트

구성 요소	버전	설명
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.206	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.206	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.1	Spark 명령줄 클라이언트.
spark-history-server	2.3.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tensorflow	1.9.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.17.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.17.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.

분류	설명
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.

분류	설명
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.17.0

5.17.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, TensorFlow, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
AWS Java용 SDK	1.11.336	1.11.336	1.11.336	1.11.333
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.2	1.5.0	1.5.0	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.6	1.4.4	1.4.4	1.4.4

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.4	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.5.0	0.4.0
MXNet	1.2.0	1.2.0	1.2.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.14.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.206	0.203	0.203	0.194
Spark	2.3.1	2.3.1	2.3.1	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	1.9.0	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.17.0	emr-5.16.1	emr-5.16.0	emr-5.15.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

5.17.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.17.0에 대한 정보가 포함됩니다. 변경 사항은 5.16.0에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 30일

업그레이드

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

새로운 특성

- Tensorflow에 대한 지원을 추가했습니다. 자세한 내용은 TensorFlow 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- JupyterHub
  - Amazon S3의 노트북 지속성에 대한 지원을 추가했습니다. 자세한 내용은 Amazon S3에서 노트북의 지속성 구성 단원을 참조하십시오.
- Spark
  - S3 Select에 대한 지원을 추가했습니다. 자세한 내용은 S3 Select와 함께 Spark를 사용하여 쿼리 성능 향상 단원을 참조하십시오.
- Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0의 CloudWatch 지표 및 자동 조정 기능 관련 문제를 해결했습니다.

알려진 문제

- Kerberos 인증을 사용하는 클러스터를 생성하고 Livy를 설치하면 단순 인증이 활성화되지 않았다는 오류가 발생하면서 Livy가 실패합니다. Livy 서버를 재부팅하면 문제가 해결됩니다. 해결 방법으로, 클러스터 생성 중에 프라이머리 노드에서 `sudo restart livy-server`를 실행하는 단계를 추가합니다.
- 생성 날짜가 2018-08-11인 Amazon Linux AMI를 기반으로 사용자 지정 Amazon Linux AMI를 사용하면 Oozie 서버가 시작에 실패합니다. Oozie를 사용하는 경우 생성 날짜가 다른 Amazon Linux AMI ID를 기반으로 사용자 지정 AMI를 생성하십시오. 다음 AWS CLI 명령을 사용하여 릴리스 날짜와 함께 2018.03 버전의 모든 HVM Amazon Linux AMIs에 대한 이미지 IDs 목록을 반환할 수 있으므로 적절한 Amazon Linux AMI를 기본으로 선택할 수 있습니다. MyRegion을 사용자의 리전 식별자(예: us-west-2)로 바꿉니다.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

5.17.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.1.3	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.5.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emr-s3-select	1.0.0	EMR S3Select 커넥터
emrfs	2.26.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.5.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.4-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.4-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.8.4-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.4-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.4-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.4-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.4-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.4-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.4-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.4-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.6	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.6	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.6	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.4.6	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.6	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.2.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.206	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.206	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.3.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tensorflow	1.9.0	TensorFlow는 고성능 수치 계산용 오픈 소스 소프트웨어 라이브러리입니다.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.17.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.17.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-s3-conf	Jupyter Notebook S3 지속성 구성
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.16.1

5.16.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, HueJupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
AWS Java용 SDK	1.11.336	1.11.336	1.11.333	1.11.333
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.5.0	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.4

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
HCatalog	2.3.3	2.3.3	2.3.3	2.3.3
Hadoop	2.8.4	2.8.4	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.3
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.2.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.5.0	0.4.0	0.4.0
MXNet	1.2.0	1.2.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	5.0.0
Phoenix	4.14.0	4.14.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.203	0.194	0.194
Spark	2.3.1	2.3.1	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.16.1	emr-5.16.0	emr-5.15.1	emr-5.15.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.12

5.16.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.16.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.1.0	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.25.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.5.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.4-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.4-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.8.4-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.4-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.4-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.4-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.4-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.4-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.4-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.4-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.4	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.4	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.4	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.4.4	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.4	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.2.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.203	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.203	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.3.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.16.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.16.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.

분류	설명
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.16.0

5.16.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, MXNet, LivyMahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
AWS Java용 SDK	1.11.336	1.11.333	1.11.333	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.5.0	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.4	1.4.2

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
HCatalog	2.3.3	2.3.3	2.3.3	2.3.2
Hadoop	2.8.4	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.3	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.2.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.5.0	0.4.0	0.4.0	0.4.0
MXNet	1.2.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	5.0.0	4.3.0
Phoenix	4.14.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.203	0.194	0.194	0.194
Spark	2.3.1	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.16.0	emr-5.15.1	emr-5.15.0	emr-5.14.2
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.12	3.4.10

5.16.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.16.0에 대한 정보가 포함됩니다. 변경 사항은 5.15.0에 관련됩니다.

최초 릴리스 날짜: 2018년 7월 19일

업그레이드

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1
- AWS SDK for Java 1.11.336
- CUDA 9.2
- Redshift JDBC 드라이버 1.2.15.1025

변경 사항, 향상된 기능 및 해결된 문제

- HBase
  - HBASE-20723 백포트
- Presto
  - LDAP 인증을 지원하도록 구성을 변경합니다. 자세한 내용은 Amazon EMR 기반 Presto에 대한 LDAP 인증 사용 단원을 참조하십시오.

- Spark
  - Amazon EMR 릴리스 5.16.0부터 사용할 수 있는 Apache Spark 버전 2.3.1은 CVE-2018-8024 및 CVE-2018-1334를 처리합니다. Spark의 이전 버전을 Spark 버전 2.3.1 이상으로 마이그레이션하는 것이 좋습니다.

알려진 문제

- 이 릴리스 버전은 c1.medium 또는 m1.small 인스턴스 유형을 지원하지 않습니다. 둘 중 어느 쪽이든 이런 인스턴스 유형을 사용하는 클러스터는 시작되지 않습니다. 차선책으로 다른 인스턴스 유형을 지정하거나 다른 릴리스 버전을 사용하십시오.
- Kerberos 인증을 사용하는 클러스터를 생성하고 Livy를 설치하면 단순 인증이 활성화되지 않았다는 오류가 발생하면서 Livy가 실패합니다. Livy 서버를 재부팅하면 문제가 해결됩니다. 해결 방법으로, 클러스터 생성 중에 프라이머리 노드에서 `sudo restart livy-server`를 실행하는 단계를 추가합니다.
- 프라이머리 노드가 재부팅되거나 인스턴스 컨트롤러가 다시 시작된 후에는 CloudWatch 지표가 수집되지 않으며 Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0에서는 자동 조정 기능을 사용할 수 없습니다. 이 문제는 Amazon EMR 5.17.0에서 수정되었습니다.

5.16.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.1.0	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.6.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.25.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.5.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.4-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.4-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.8.4-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.4-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.4-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.4-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.4-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.4-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.4-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.4-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.4	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.4	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.4	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.4.4	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.4	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.5.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.2.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.2.88	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.14.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.14.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.203	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.203	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.3.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.16.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.16.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.

분류	설명
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-password-authenticator	Presto의 password-authenticator.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.15.1

5.15.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
AWS Java용 SDK	1.11.333	1.11.333	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.4	1.4.2	1.4.2

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
HCatalog	2.3.3	2.3.3	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.3	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.2.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	5.0.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.15.1	emr-5.15.0	emr-5.14.2	emr-5.14.1
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.12	3.4.10	3.4.10

5.15.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.15.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK

구성 요소	버전	설명
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.24.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.8.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.4	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.4	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.4	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.4.4	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.4	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.3-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.1.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.194	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.194	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.3.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.15.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.15.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.

분류	설명
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.15.0

5.15.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
AWS Java용 SDK	1.11.333	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.4	1.4.2	1.4.2	1.4.2

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
HCatalog	2.3.3	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.3	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.2.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	0.8.1
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.1.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	5.0.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.7
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.15.0	emr-5.14.2	emr-5.14.1	emr-5.14.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.12	3.4.10	3.4.10	3.4.10

5.15.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.15.0에 대한 정보가 포함됩니다. 변경 사항은 5.14.0에 관련됩니다.

최초 릴리스 날짜: 2018년 6월 21일

업그레이드

- HBase를 1.4.4로 업그레이드
- Hive를 2.3.3으로 업그레이드
- Hue를 4.2.0으로 업그레이드
- Oozie를 5.0.0로 업그레이드
- Zookeeper를 3.4.12로 업그레이드
- AWS SDK를 1.11.333으로 업그레이드

변경 사항, 향상된 기능 및 해결된 문제

- Hive
  - HIVE-18069를 백포트했습니다.
- Hue
  - Kerberos 사용 시 Livy로 올바로 인증하도록 Hue를 업데이트했습니다. 이제 Livy가 Amazon EMR에서 Kerberos를 사용할 때 지원됩니다.
- JupyterHub
  - Amazon EMR이 LDAP 클라이언트 라이브러리를 기본으로 설치하도록 JupyterHub를 업데이트했습니다.
  - 자체 서명 인증서를 생성하는 스크립트 오류를 수정했습니다.

알려진 문제

- 이 릴리스 버전은 c1.medium 또는 m1.small 인스턴스 유형을 지원하지 않습니다. 둘 중 어느 쪽이든 이런 인스턴스 유형을 사용하는 클러스터는 시작되지 않습니다. 차선책으로 다른 인스턴스 유형을 지정하거나 다른 릴리스 버전을 사용하십시오.
- 프라이머리 노드가 재부팅되거나 인스턴스 컨트롤러가 다시 시작된 후에는 CloudWatch 지표가 수집되지 않으며 Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0에서는 자동 조정 기능을 사용할 수 없습니다. 이 문제는 Amazon EMR 5.17.0에서 수정되었습니다.

5.15.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.24.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.8.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.4	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.4	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.4	HBase 명령줄 클라이언트
hbase-rest-server	1.4.4	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.4	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.3-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	2.3.3-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	2.3.3-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.3-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.3-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.3-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.3-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.2.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.1.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트

구성 요소	버전	설명
oozie-client	5.0.0	Oozie 명령줄 클라이언트
oozie-server	5.0.0	Oozie 워크플로 요청을 수락하는 서비스
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.194	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.194	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.0	Spark 명령줄 클라이언트.
spark-history-server	2.3.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.

구성 요소	버전	설명
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.12	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.12	ZooKeeper 명령줄 클라이언트.

5.15.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-5.15.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.

분류	설명
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.14.2

5.14.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
AWS Java용 SDK	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.2	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-

	emr-5.14.2	emr-5.14.1	emr-5.14.0	emr-5.13.1
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	0.8.1	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.1.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.7	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.14.2 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.14.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.23.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.8.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.2	HBase 명령줄 클라이언트
hbase-rest-server	1.4.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	2.3.2-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.2-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.1.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트

구성 요소	버전	설명
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.194	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.194	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.0	Spark 명령줄 클라이언트.
spark-history-server	2.3.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.

구성 요소	버전	설명
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.14.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.14.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.

분류	설명
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.14.1

5.14.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
AWS Java용 SDK	1.11.297	1.11.297	1.11.297	1.11.297
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.2	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.2
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-

	emr-5.14.1	emr-5.14.0	emr-5.13.1	emr-5.13.0
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	0.8.1	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.1.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.194
Spark	2.3.0	2.3.0	2.3.0	2.3.0
Sqoop	1.4.7	1.4.7	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.14.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.14.1에 대한 정보가 포함됩니다. 변경 사항은 5.14.0에 관련됩니다.

최초 릴리스 날짜: 2018년 10월 17일

잠재적 보안 취약성을 해결하기 위해 Amazon EMR의 기본 AMI를 업데이트했습니다.

5.14.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.23.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.8.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.2	HBase 명령줄 클라이언트
hbase-rest-server	1.4.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.2-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.2-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.1.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.

구성 요소	버전	설명
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.194	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.194	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.0	Spark 명령줄 클라이언트.
spark-history-server	2.3.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.14.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.14.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.

분류	설명
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.

분류	설명
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.

분류	설명
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.14.0

5.14.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, JupyterHub, Livy, MXNet, MahoutOozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
AWS Java용 SDK	1.11.297	1.11.297	1.11.297	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.2	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.2	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	0.8.1	-	-	-

	emr-5.14.0	emr-5.13.1	emr-5.13.0	emr-5.12.3
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.1.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.194	0.188
Spark	2.3.0	2.3.0	2.3.0	2.2.1
Sqoop	1.4.7	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.14.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.14.0에 대한 정보가 포함됩니다. 변경 사항은 5.13.0에 관련됩니다.

최초 릴리스 날짜: 2018년 6월 4일

업그레이드

- Apache Flink를 1.4.2로 업그레이드

- Apache MXnet을 1.1.0으로 업그레이드
- Apache Sqoop를 1.4.7로 업그레이드

새로운 특성

- JupyterHub 지원을 추가했습니다. 자세한 내용은 JupyterHub 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- EMRFS
  - Amazon S3에 대한 요청의 userAgent 문자열이 간접 호출하는 보안 주체의 사용자 및 그룹 정보를 포함하도록 업데이트되었습니다. 이는 보다 포괄적인 요청 추적을 위해 AWS CloudTrail 로그와 함께 사용할 수 있습니다.
- HBase
  - 포함된 HBASE-20447은 특히 분할된 리전과 함께 캐시 문제를 야기할 수 있는 문제를 해결합니다.
- MXnet
  - OpenCV 라이브러리를 추가했습니다.
- Spark
  - Spark가 EMRFS를 사용하여 Amazon S3 위치에 Parquet 파일을 쓸 때, FileOutputCommitter 알고리즘이 버전 1 대신 버전 2를 사용하도록 업데이트되었습니다. 따라서 이름 변경 횟수가 감소하여 애플리케이션 성능이 개선됩니다. 다음 애플리케이션에는 이 변경이 영향을 미치지 않습니다.
    - Spark 이외의 애플리케이션
    - HDFS(여전히 버전 1의 FileOutputCommitter를 사용)와 같이 다른 파일 시스템에 쓰는 애플리케이션
    - 텍스트 또는 csv 같이 이미 EMRFS 직접 쓰기를 사용하는 다른 출력 형식을 사용하는 애플리케이션

알려진 문제

- JupyterHub
  - 클러스터를 생성할 때 구성 분류를 사용하여 JupyterHub 및 개별 Jupyter Notebook을 설정하는 것은 지원되지 않습니다. 각 사용자에 대해 수동으로 jupyterhub_config.py 파일 및

jupyter_notebook_config.py 파일을 편집하십시오. 자세한 내용은 JupyterHub 구성 단원을 참조하십시오.

- JupyterHub는 프라이빗 서브넷 내의 클러스터에서 시작되지 않으며 `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` 메시지와 함께 실패합니다. 이는 자체 서명된 인증서를 생성하는 스크립트의 오류에 의해 야기됩니다. 다음 차선책을 사용하여 자체 서명된 인증서를 생성합니다. 모든 명령은 프라이머리 노드에 연결되어 있는 동안 실행됩니다.

  1. 컨테이너에서 프라이머리 노드로 인증서 생성 스크립트를 복사합니다.

     ```
     sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
     ```

  2. 아래와 같이 텍스트 편집기를 사용하여 23행을 변경하고 퍼블릭 호스트 이름을 로컬 호스트 이름으로 변경합니다.

     ```
     local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
     ```

  3. 스크립트를 실행하여 자체 서명된 인증서를 생성합니다.

     ```
     sudo bash ./gen_self_signed_cert.sh
     ```

  4. 스크립트가 생성하는 인증서 파일을 `/etc/jupyter/conf/` 디렉터리로 이동합니다.

     ```
     sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
     ```

  JupyterHub의 재시작과 200 응답 코드의 반환을 확인하기 위해 `jupyter.log` 파일을 `tail`할 수 있습니다. 예시:

  ```
  tail -f /var/log/jupyter/jupyter.log
  ```

  그러면 다음과 비슷한 응답이 돌아옵니다.

  ```
  # [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
   https://:9443/
  # 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
  ```

- 프라이머리 노드가 재부팅되거나 인스턴스 컨트롤러가 다시 시작된 후에는 CloudWatch 지표가 수집되지 않으며 Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0에서는 자동 조정 기능을 사용할 수 없습니다. 이 문제는 Amazon EMR 5.17.0에서 수정되었습니다.

5.14.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.23.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.8.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.2	HBase 명령줄 클라이언트
hbase-rest-server	1.4.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.2-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-2	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.3.2-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
jupyterhub	0.8.1	Jupyter Notebook용 다중 사용자 서버
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.1.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
opencv	3.4.0	오픈 소스 컴퓨터 비전 라이브러리.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.194	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.194	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.0	Spark 명령줄 클라이언트.
spark-history-server	2.3.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.7	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.14.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.14.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
container-log4j	하둡 YARN의 container-log4j.properties 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
jupyter-notebook-conf	Jupyter Notebook의 jupyter_notebook_config.py 파일에서 값을 변경합니다.
jupyter-hub-conf	JupyterHubs의 jupyterhub_config.py 파일에서 값을 변경합니다.
jupyter-sparkmagic-conf	Sparkmagic의 config.json 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.

분류	설명
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.13.1

5.13.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
AWS Java용 SDK	1.11.297	1.11.297	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.2	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0

	emr-5.13.1	emr-5.13.0	emr-5.12.3	emr-5.12.2
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.194	0.188	0.188
Spark	2.3.0	2.3.0	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.13.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.13.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.22.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하듭 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하듭 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-0	하듭의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.8.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.2	HBase 명령줄 클라이언트
hbase-rest-server	1.4.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.2-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-2	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.3.2-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.0.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.194	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.194	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.0	Spark 명령줄 클라이언트.
spark-history-server	2.3.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.

구성 요소	버전	설명
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.13.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.13.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.

분류	설명
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.

분류	설명
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.

분류	설명
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.13.0

5.13.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, MXNet, LivyMahout, Oozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
AWS Java용 SDK	1.11.297	1.11.267	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.2	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.194	0.188	0.188	0.188
Spark	2.3.0	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-5.13.0	emr-5.12.3	emr-5.12.2	emr-5.12.1
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.13.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.13.0에 대한 정보가 포함됩니다. 변경 사항은 5.12.0에 관련됩니다.

업그레이드

- Spark를 2.3.0으로 업그레이드
- HBase를 1.4.2로 업그레이드
- Presto를 0.194로 업그레이드
- 1.11.297 AWS SDK for Java 로 업그레이드됨

변경 사항, 향상된 기능 및 해결된 문제

- Hive
  - HIVE-15436을 백포트했습니다. Hive API를 뷰만 반환하도록 개선했습니다.

알려진 문제

- 현재 MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.

5.13.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.10.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.22.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.8.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.4.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.2	HBase 명령줄 클라이언트
hbase-rest-server	1.4.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-2	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.2-amzn-2	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-2	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-2	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-2	Hive-hbase 클라이언트
hive-metastore-server	2.3.2-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.0.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.194	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.194	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
r	3.4.1	통계 컴퓨팅용 R 프로젝트
spark-client	2.3.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.3.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.3.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.3.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.13.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.13.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.12.3

5.12.3 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
AWS Java용 SDK	1.11.267	1.11.267	1.11.267	1.11.267
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.4.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.4.0
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.8.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.1.0	4.1.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.12.3	emr-5.12.2	emr-5.12.1	emr-5.12.0
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	1.0.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.13.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.188
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.12.3 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.12.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.9.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.21.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.8.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.0	HBase 명령줄 클라이언트
hbase-rest-server	1.4.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.2-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-1	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.3.2-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.0.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.188	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.188	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.12.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.12.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.12.2

5.12.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
AWS Java용 SDK	1.11.267	1.11.267	1.11.267	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.4.0	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.4.0	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.8.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-

	emr-5.12.2	emr-5.12.1	emr-5.12.0	emr-5.11.4
Hue	4.1.0	4.1.0	4.1.0	4.0.1
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	1.0.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.13.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.188	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.12.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.12.2에 대한 정보가 포함됩니다. 변경 사항은 5.12.1에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

5.12.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.9.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.21.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.8.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.0	HBase 명령줄 클라이언트
hbase-rest-server	1.4.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	2.3.2-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.2-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.0.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.

구성 요소	버전	설명
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.188	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.188	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.12.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.12.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.

분류	설명
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.12.1

5.12.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
AWS Java용 SDK	1.11.267	1.11.267	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
Delta	-	-	-	-
Flink	1.4.0	1.4.0	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.4.0	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.8.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.1.0	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	1.0.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.13.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.188	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1

	emr-5.12.1	emr-5.12.0	emr-5.11.4	emr-5.11.3
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.12.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.12.1에 대한 정보가 포함됩니다. 변경 사항은 5.12.0에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

5.12.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른

Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.9.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.21.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.8.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.8.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.4.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.0	HBase 명령줄 클라이언트
hbase-rest-server	1.4.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.3.2-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.2-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.2-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.0.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.188	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.188	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.12.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.12.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.12.0

5.12.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
AWS Java용 SDK	1.11.267	1.11.238	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.4.0	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.4.0	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.8.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.1.0	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.12.0	emr-5.11.4	emr-5.11.3	emr-5.11.2
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	1.0.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.13.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.188	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.12.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.12.0에 대한 정보가 포함됩니다. 변경 사항은 5.11.1에 관련됩니다.

업그레이드

- AWS SDK for Java 1.11.238 " 1.11.267. 자세한 내용은 GitHub에서 AWS SDK for Java Change Log를 참조하세요.
- Hadoop 2.7.3 ⇒ 2.8.3. 자세한 내용은 Apache Hadoop Releases를 참조하십시오.

- Flink 1.3.2 ⇒ 1.4.0. 자세한 내용은 Apache Flink 1.4.0 Release Announcement를 참조하십시오.
- HBase 1.3.1 ⇒ 1.4.0. 자세한 내용은 HBase Release Announcement를 참조하십시오.
- Hue 4.0.1 ⇒ 4.1.0. 자세한 내용은 릴리스 정보를 참조하십시오.
- MxNet 0.12.0 ⇒ 1.0.0. 자세한 내용은 GitHub에서 MXNet Change Log를 참조하십시오.
- Presto 0.187 ⇒ 0.188. 자세한 내용은 릴리스 정보를 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- Hadoop
  - `yarn.resourcemanager.decommissioning.timeout` 속성을 `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`로 변경했습니다. 이 속성을 사용하여 클러스터 축소를 사용자 지정할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 스케일 다운을 참조하세요.
  - Hadoop CLI는 직접 복사를 지정하는 `-d`(복사) 명령에 `cp` 옵션을 추가했습니다. 이를 사용하면 중간 `.COPYING` 파일을 생성하지 않아도 되므로 Amazon S3 사이에서 데이터를 더 빠르게 복사할 수 있습니다. 자세한 내용은 HADOOP-12384를 참조하십시오.
- Pig
  - Pig 환경 속성의 구성을 단순화하는 `pig-env` 구성 분류를 추가했습니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.
- Presto
  - Presto `presto-connector-redshift` 구성 파일에서 값을 구성하는 데 사용할 수 있는 `redshift.properties` 구성 분류를 추가했습니다. 자세한 내용은 Presto 문서의 Redshift Connector 및 애플리케이션 구성 단원을 참조하십시오.
  - EMRFS에 대한 Presto 지원이 추가되었으며 이것이 기본 구성입니다. 이전 Amazon EMR 릴리스에서는 유일한 옵션인 PrestoS3FileSystem을 사용했습니다. 자세한 내용은 EMRFS 및 PrestoS3FileSystem 구성 단원을 참조하십시오.

> **Note**
>
> Amazon EMR 버전 5.12.0을 사용하여 Amazon S3에서 기본 데이터를 쿼리하는 경우 Presto 오류가 발생할 수 있습니다. 이는 Presto가 `emrfs-site.xml`에서 구성 분류 값을 가져오지 못하기 때문입니다. 이 문제를 해결하려면 `usr/lib/presto/plugin/hive-hadoop2/` 아래에 `emrfs` 하위 디렉터리를 생성하고 `usr/lib/presto/plugin/hive-hadoop2/emrfs`에서 기존 `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` 파일에 대한 기호 링크를 생성합니다. 그런 다음, presto-server 프로세스를 다

> 시 시작합니다(`sudo presto-server stop` 및 `sudo presto-server start`의 순서대로)

- Spark
  - SPARK-22036: BigDecimal multiplication sometimes returns null을 백포트했습니다.

알려진 문제

- MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.
- 클러스터 노드에 R이 기본적으로 설치되어 있지 않으므로 사용자 지정 AMI를 사용하여 만든 클러스터에는 SparkR을 사용할 수 없습니다.

5.12.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0.1	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.

구성 요소	버전	설명
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.9.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.21.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.4.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.8.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.8.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.8.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.8.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	2.8.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.8.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.8.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.8.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.8.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.8.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.4.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.4.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.4.0	HBase 명령줄 클라이언트
hbase-rest-server	1.4.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.4.0	HBase에 Thrift 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hcatalog-client	2.3.2-amzn-1	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.2-amzn-1	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-1	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-1	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-1	Hive-hbase 클라이언트
hive-metastore-server	2.3.2-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.1.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	1.0.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.1.85	Nvidia 드라이버와 Cuda 도구 키트

구성 요소	버전	설명
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.13.0-HBase-1.4	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.13.0-HBase-1.4	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.188	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.188	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.12.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.12.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-env	Pig 환경에서 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.

분류	설명
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-redshift	Presto redshift.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.11.4

5.11.4 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
AWS Java용 SDK	1.11.238	1.11.238	1.11.238	1.11.238
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1

	emr-5.11.4	emr-5.11.3	emr-5.11.2	emr-5.11.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.11.4 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.11.4 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.8.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.20.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.7.3-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리.
hadoop-hdfs-namenode	2.7.3-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트.
hadoop-kms-server	2.7.3-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버.
hadoop-mapred	2.7.3-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리.
hadoop-yarn-nodemanager	2.7.3-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스.
hadoop-yarn-resourcemanager	2.7.3-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스.
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스.
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.

구성 요소	버전	설명
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.2-amzn-0	hcatalog-server에 대하 작업을 수행하기 위한 'hcat' 명령줄 클라이언트.
hcatalog-server	2.3.2-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스.
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스

구성 요소	버전	설명
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	0.12.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.0.176	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.187	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.187	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.

구성 요소	버전	설명
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.11.4 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.11.4 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.

분류	설명
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.

분류	설명
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.11.3

5.11.3 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
AWS Java용 SDK	1.11.238	1.11.238	1.11.238	1.11.238

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.2
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.11.3	emr-5.11.2	emr-5.11.1	emr-5.11.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.11.3 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.3에 대한 정보가 포함됩니다. 변경 사항은 5.11.2에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 18일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

5.11.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.8.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.20.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리.
hadoop-hdfs-namenode	2.7.3-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트.
hadoop-kms-server	2.7.3-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버.
hadoop-mapred	2.7.3-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리.
hadoop-yarn-nodemanager	2.7.3-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스.
hadoop-yarn-resourcemanager	2.7.3-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스.

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스.
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.2-amzn-0	hcatalog-server에 대하 작업을 수행하기 위한 'hcat' 명령줄 클라이언트.
hcatalog-server	2.3.2-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.3.2-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스.
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	0.12.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.0.176	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.187	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.187	여러 쿼리를 실행하는 서비스.

구성 요소	버전	설명
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.11.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.11.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.

분류	설명
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.11.2

5.11.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
AWS Java용 SDK	1.11.238	1.11.238	1.11.238	1.11.221
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.2	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.2	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0

	emr-5.11.2	emr-5.11.1	emr-5.11.0	emr-5.10.1
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.1	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.11.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.2에 대한 정보가 포함됩니다. 변경 사항은 5.11.1에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

5.11.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.8.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.20.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리.
hadoop-hdfs-namenode	2.7.3-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트.
hadoop-kms-server	2.7.3-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버.
hadoop-mapred	2.7.3-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리.
hadoop-yarn-nodemanager	2.7.3-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스.

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.7.3-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스.
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스.
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.2-amzn-0	hcatalog-server에 대하 작업을 수행하기 위한 'hcat' 명령줄 클라이언트.
hcatalog-server	2.3.2-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-0	Hive-hbase client.

구성 요소	버전	설명
hive-metastore-server	2.3.2-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스.
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	0.12.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.0.176	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.187	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.187	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.11.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.11.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.11.1

5.11.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
AWS Java용 SDK	1.11.238	1.11.238	1.11.221	1.11.221
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	2.7, 3.4
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.2	2.3.1	2.3.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.2	2.3.1	2.3.1
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-5.11.1	emr-5.11.0	emr-5.10.1	emr-5.10.0
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	0.12.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.187
Spark	2.2.1	2.2.1	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.3
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.11.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.8.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

5.11.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.8.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.20.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리.
hadoop-hdfs-namenode	2.7.3-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트.
hadoop-kms-server	2.7.3-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버.

구성 요소	버전	설명
hadoop-mapred	2.7.3-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리.
hadoop-yarn-nodemanager	2.7.3-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스.
hadoop-yarn-resourcemanager	2.7.3-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스.
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스.
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.2-amzn-0	hcatalog-server에 대하 작업을 수행하기 위한 'hcat' 명령줄 클라이언트.
hcatalog-server	2.3.2-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스.
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	0.12.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.0.176	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.

구성 요소	버전	설명
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.187	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.187	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.11.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.11.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.11.0

5.11.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Livy, Hue, Mahout, MXNetOozie, Phoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
AWS Java용 SDK	1.11.238	1.11.221	1.11.221	1.11.183
Python	2.7, 3.4	2.7, 3.4	2.7, 3.4	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.2	2.3.1	2.3.1	2.3.0

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.2	2.3.1	2.3.1	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	0.12.0	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.187	0.184
Spark	2.2.1	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-5.11.0	emr-5.10.1	emr-5.10.0	emr-5.9.1
Zeppelin	0.7.3	0.7.3	0.7.3	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.11.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.0에 대한 정보가 포함됩니다. 변경 사항은 5.10.0에 관련됩니다.

업그레이드

- Hive 2.3.2
- Spark 2.2.1
- SDK for Java 1.11.238

새로운 특성

- Spark
  - 스팟 인스턴스 사용 시 Spark 폐기를 개선하는 `spark.decommissioning.timeout.threshold` 설정이 추가되었습니다. 자세한 내용은 노드 서비스 해제 동작 구성 단원을 참조하십시오.
  - Amazon SageMaker Spark 및 Amazon SageMaker와 Spark의 통합을 위한 관련 종속 항목을 설치하는 `aws-sagemaker-spark-sdk` 구성 요소가 Spark에 추가되었습니다. Amazon SageMaker Spark를 사용하면 Amazon SageMaker 단계를 사용하여 Spark 기계 학습(ML) 파이프라인을 생성할 수 있습니다. 자세한 내용은 GitHub의 SageMaker Spark Readme와 Amazon SageMaker 개발자 안내서의 Amazon SageMaker와 함께 Apache Spark 사용을 참조하십시오.

알려진 문제

- MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.
- Hive 2.3.2에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을

LOCATION 테이블에 업로드하고 테이블에서 SELECT COUNT(*) 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

차선책으로 ANALYZE TABLE 명령을 사용하여 새 통계를 수집하거나 hive.compute.query.using.stats=false를 설정합니다. 자세한 내용은 Apache Hive 설명서의 Statistics in Hive를 참조하십시오.

5.11.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 emr 또는 aws로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 myapp-component라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 2.2-amzn-2로 나열됩니다.

구성 요소	버전	설명
aws-sagemaker-spark-sdk	1.0	Amazon SageMaker Spark SDK
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.8.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.20.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-6	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-6	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-6	HDFS 명령줄 클라이언트 및 라이브러리.
hadoop-hdfs-namenode	2.7.3-amzn-6	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-6	HDFS 작업에 대한 HTTP 엔드포인트.
hadoop-kms-server	2.7.3-amzn-6	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버.

구성 요소	버전	설명
hadoop-mapred	2.7.3-amzn-6	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리.
hadoop-yarn-nodemanager	2.7.3-amzn-6	개별 노드의 컨테이너를 관리하는 YARN 서비스.
hadoop-yarn-resourcemanager	2.7.3-amzn-6	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스.
hadoop-yarn-timeline-server	2.7.3-amzn-6	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스.
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.2-amzn-0	hcatalog-server에 대하 작업을 수행하기 위한 'hcat' 명령줄 클라이언트.
hcatalog-server	2.3.2-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.3.2-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.2-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.2-amzn-0	Hive-hbase client.
hive-metastore-server	2.3.2-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.2-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스.
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	0.12.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.0.176	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.

구성 요소	버전	설명
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.187	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.187	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.1	Spark 명령줄 클라이언트.
spark-history-server	2.2.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.11.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.11.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.10.1

5.10.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
AWS Java용 SDK	1.11.221	1.11.221	1.11.183	1.11.183
Python	2.7, 3.4	2.7, 3.4	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.2
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
HCatalog	2.3.1	2.3.1	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.1	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	4.0.1
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	0.4.0
MXNet	0.12.0	0.12.0	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.17.0
Presto	0.187	0.187	0.184	0.184
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4

	emr-5.10.1	emr-5.10.0	emr-5.9.1	emr-5.9.0
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.3	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.10.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.10.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.7.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.20.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트.
hadoop-hdfs-datanode	2.7.3-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스.

구성 요소	버전	설명
hadoop-hdfs-library	2.7.3-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스.
hadoop-httpfs-server	2.7.3-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.

구성 요소	버전	설명
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층 및 테이블, HCatalog를 제공하는 서비스.
hcatalog-webhcat-server	2.3.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트.
hive-client	2.3.1-amzn-0	Hive 명령줄 클라이언트.
hive-hbase	2.3.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스.
hive-server2	2.3.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.

구성 요소	버전	설명
mxnet	0.12.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.0.176	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.187	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.187	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.

구성 요소	버전	설명
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.10.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.10.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.

분류	설명
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.

분류	설명
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.10.0

5.10.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, Hive, Hue, Livy, MXNet, Mahout, OoziePhoenix, Pig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
AWS Java용 SDK	1.11.221	1.11.183	1.11.183	1.11.160

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
Python	2.7, 3.4	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.2	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.1	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.1	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	4.0.1	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	0.4.0	-
MXNet	0.12.0	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0

	emr-5.10.0	emr-5.9.1	emr-5.9.0	emr-5.8.3
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.17.0	0.16.0
Presto	0.187	0.184	0.184	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.3	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.10.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.10.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.9.0 릴리스에 관련됩니다.

업그레이드

- AWS SDK for Java 1.11.221
- Hive 2.3.1
- Presto 0.187

새로운 특성

- Kerberos 인증에 대한 지원이 추가되었습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Kerberos 인증 사용을 참조하세요.

- EMRFS의 IAM 역할에 대한 지원이 추가되었습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon S3에 대한 EMRFS 요청을 위해 IAM 역할 구성을 참조하세요.
- GPU 기반의 P2 및 P3 인스턴스 유형에 대한 지원을 추가했습니다. 자세한 내용은 Amazon EC2 P2 인스턴스 및 Amazon EC2 P3 인스턴스를 참조하세요. 기본적으로 이러한 유형의 인스턴스에 NVIDIA 드라이버 384.81 및 CUDA 드라이버 9.0.176가 설치됩니다.
- Apache MXNet 지원이 추가되었습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Presto
  - Glue 데이터 카탈로그를 기본 Hive AWS 메타스토어로 사용하기 위한 지원이 추가되었습니다. 자세한 내용은 AWS Glue 데이터 카탈로그에서 Presto 사용을 참조하세요.
  - 지리 공간 함수에 대한 지원이 추가되었습니다.
  - 조인의 디스크로 유출에 대한 지원이 추가되었습니다.
  - Redshift 커넥터에 대한 지원이 추가되었습니다.
- Spark
  - SPARK-20640을 백포트했습니다. 이를 통해 `spark.shuffle.registration.timeout` 및 `spark.shuffle.registration.maxAttempts` 속성을 사용하여 셔플 등록에 대한 제시도와 rpc 제한 시간을 구성할 수 있습니다.
  - 사용자 지정 OutputFormat를 비 HDFS 위치에 기록할 때 발생하는 오류를 교정하는 SPARK-21549를 백포트했습니다.
- 하둡-13270을 백포트했습니다.
- Numpy, Scipy 및 Matplotlib 라이브러리가 기본 Amazon EMR AMI에서 제거되었습니다. 애플리케이션에서 이러한 라이브러리가 필요한 경우, 애플리케이션 리포지토리에서 라이브러리를 사용할 수 있기 때문에 부트스트랩 작업 시 `yum install`을 사용하여 모든 노드에 라이브러리를 설치할 수 있습니다.
- Amazon EMR 기본 AMI는 더 이상 애플리케이션 RPM 패키지를 포함하고 있지 않기 때문에 RPM 패키지가 더 이상 클러스터 노드에 존재하지 않습니다. 이제 사용자 지정 AMI 및 Amazon EMR 기본 AMI는 Amazon S3에서 RPM 패키지 리포지토리를 참조합니다.
- Amazon EC2에 초당 요금이 도입됨에 따라 이제 기본 스케일 다운 동작은 인스턴스 시간에 종료가 아닌 작업 완료 시 종료입니다. 자세한 내용은 클러스터 축소 구성을 참조하십시오.

알려진 문제

- MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.
- Hive 2.3.1에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 Statistics in Hive를 참조하십시오.

5.10.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.5.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터.
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.7.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.20.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-5	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트.
hadoop-hdfs-datanode	2.7.3-amzn-5	블록을 저장하는 HDFS 노드 수준 서비스.
hadoop-hdfs-library	2.7.3-amzn-5	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-5	파일 이름 및 블록 위치를 추적하는 HDFS 서비스.
hadoop-httpfs-server	2.7.3-amzn-5	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.7.3-amzn-5	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-5	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-5	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-5	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-5	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	2.3.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층 및 테이블, HCatalog 를 제공하는 서비스.
hcatalog-webhcat-server	2.3.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트.
hive-client	2.3.1-amzn-0	Hive 명령줄 클라이언트.
hive-hbase	2.3.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스.
hive-server2	2.3.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mxnet	0.12.0	유연성, 확장성 및 효율성이 뛰어난 딥 러닝용 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
nvidia-cuda	9.0.176	Nvidia 드라이버와 Cuda 도구 키트
oozie-client	4.3.0	Oozie 명령줄 클라이언트.

구성 요소	버전	설명
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.187	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.187	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.7.3	대화형 데이터 분석이 가능한 웹 기반 노트북.
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.10.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.10.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.

분류	설명
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.9.1

5.9.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, LivyOozie, PhoenixPig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
AWS Java용 SDK	1.11.183	1.11.183	1.11.160	1.11.160
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.2	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	4.0.1	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	0.4.0	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.17.0	0.16.0	0.16.0
Presto	0.184	0.184	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-5.9.1	emr-5.9.0	emr-5.8.3	emr-5.8.2
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.9.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.9.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.4.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.

구성 요소	버전	설명
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.7.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.19.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	2.7.3-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.

구성 요소	버전	설명
hcatalog-client	2.3.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.0-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.0-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.

구성 요소	버전	설명
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.184	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.184	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.9.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.9.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.

분류	설명
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.9.0

5.9.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, LivyMahout, Oozie, PhoenixPig, Presto, Spark, Sqoop, Tez, 및 애플리케이션이 포함되어 Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
AWS Java용 SDK	1.11.183	1.11.160	1.11.160	1.11.160
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.2	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	4.0.1	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	0.4.0	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.17.0	0.16.0	0.16.0	0.16.0
Presto	0.184	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-5.9.0	emr-5.8.3	emr-5.8.2	emr-5.8.1
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.9.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.9.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.8.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 10월 5일

최근 기능 업데이트: 2017년 10월 12일

업그레이드

- AWS SDK for Java 버전 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

새로운 특성

- Livy 지원이 추가되었습니다(버전 0.4.0-incubating). 자세한 내용은 Apache Livy 단원을 참조하십시오.
- Spark용 Hue Notebook에 대한 지원이 추가되었습니다.
- i3 시리즈 Amazon EC2 인스턴스에 대한 지원이 추가되었습니다(2017년 10월 12일).

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Spark가 수동 크기 조절 또는 자동 조정 정책 요청으로 인한 노드 종료를 보다 원활하게 처리할 수 있도록 도움을 주는 새로운 기능 세트를 추가했습니다. 자세한 내용은 노드 서비스 해제 동작 구성 단원을 참조하십시오.

  - SSL은 블록 전송 서비스의 전송 중 암호화 시 3DES를 대신해 사용되며, AES-NI와 함께 Amazon EC2 인스턴스 유형을 사용할 때 성능을 개선합니다.
  - SPARK-21494를 백포트했습니다.
- Zeppelin
  - ZEPPELIN-2377을 백포트했습니다.
- HBase
  - HBase-18533 패치를 추가했습니다. 이 패치는 `hbase-site` 구성 분류를 사용하여 HBase BucketCache 구성을 위한 추가적인 값들을 허용합니다.
- Hue
  - Hue의 Hive 쿼리 편집기에 대한 AWS Glue 데이터 카탈로그 지원이 추가되었습니다.
  - 기본적으로 Hue의 수퍼유저는 Amazon EMR IAM 역할에서 액세스할 수 있는 모든 파일에 액세스할 수 있습니다. 새로 생성된 사용자에게는 Amazon S3 파일 브라우저에 액세스할 수 있는 권한이 자동으로 부여되지 않기 때문에 사용자의 그룹에 대해 활성화된 `filebrowser.s3_access` 권한이 있어야 합니다.
- AWS Glue 데이터 카탈로그를 사용하여 생성된 기본 JSON 데이터에 액세스할 수 없는 문제를 해결했습니다.

알려진 문제

- 모든 애플리케이션을 설치하고 기본 Amazon EBS 루트 볼륨 크기를 변경하지 않은 상태에서 클러스터가 시작되지 않습니다. 해결 방법으로의 `aws emr create-cluster` 명령을 사용하고 더 큰 `--ebs-root-volume-size` 파라미터를 AWS CLI 지정합니다.
- Hive 2.3.0에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 Statistics in Hive를 참조하십시오.

5.9.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.4.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.7.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.19.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.2	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션입니다.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.7.3-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.0-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.0-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.3.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	4.0.1	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
livy-server	0.4.0-incubating	Apache Spark와 상호작용하기 위한 REST 인터페이스
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.184	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.184	여러 쿼리를 실행하는 서비스.
pig-client	0.17.0	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.9.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.9.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
livy-conf	Livy의 livy.conf 파일의 값을 변경합니다.
livy-env	Livy 환경의 값을 변경합니다.
livy-log4j	Livy log4j.properties 설정을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.

분류	설명
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.8.3

5.8.3 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
AWS Java용 SDK	1.11.160	1.11.160	1.11.160	1.11.160
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.3.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.3.0
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.8.3	emr-5.8.2	emr-5.8.1	emr-5.8.0
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.2.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.8.3 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.8.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.4.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.6.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.18.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.0-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.0-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.3.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-1	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.

구성 요소	버전	설명
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.8.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.8.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.

분류	설명
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.

분류	설명
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.

분류	설명
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.

분류	설명
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.8.2

5.8.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Mahout, HueOozie, PhoenixPig, Presto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
AWS Java용 SDK	1.11.160	1.11.160	1.11.160	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.1	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.3.0	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.3.0	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.8.2	emr-5.8.1	emr-5.8.0	emr-5.7.1
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.2.0	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.8.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.8.2에 대한 정보가 포함됩니다. 변경 사항은 5.8.1에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

5.8.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 myapp-component라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 2.2-amzn-2로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.4.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.6.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.18.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.7.3-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.

구성 요소	버전	설명
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.0-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.0-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.

구성 요소	버전	설명
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-1	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.8.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.8.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.

분류	설명
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.

분류	설명
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.8.1

5.8.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, MahoutOozie, Pig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
AWS Java용 SDK	1.11.160	1.11.160	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.1	1.3.0	1.3.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.1
HCatalog	2.3.0	2.3.0	2.1.1	2.1.1

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.3.0	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.11.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.2.0	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-5.8.1	emr-5.8.0	emr-5.7.1	emr-5.7.0
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.2
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.8.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.8.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.8.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

5.8.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.4.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터

구성 요소	버전	설명
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.6.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.18.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리

구성 요소	버전	설명
hadoop-hdfs-namenode	2.7.3-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.

구성 요소	버전	설명
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.0-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.0-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.

구성 요소	버전	설명
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-1	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.8.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.8.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.8.0

5.8.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Mahout, HueOozie, PhoenixPig, Presto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
AWS Java용 SDK	1.11.160	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.1	1.3.0	1.3.0	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.1	1.3.0
HCatalog	2.3.0	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.3.0	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.8.0	emr-5.7.1	emr-5.7.0	emr-5.6.1
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.11.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.2.0	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	0.7.2	0.7.2	0.7.2	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.8.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.8.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.7.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2017년 8월 10일

최근 기능 업데이트: 2017년 9월 25일

업그레이드

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0 자세한 내용은 Apache Hive 사이트의 릴리스 정보를 참조하십시오.
- Spark 2.2.0 자세한 내용은 Apache Spark 사이트의 릴리스 정보를 참조하십시오.

새로운 특성

- 애플리케이션 이력 보기 지원이 추가되었습니다(2017년 9월 25일). 자세한 내용은 Amazon EMR 관리 안내서에서 애플리케이션 기록 보기를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- AWS Glue 데이터 카탈로그와 통합
  - Hive 및 Spark SQL이 AWS Glue 데이터 카탈로그를 Hive 메타데이터 스토어로 사용할 수 있는 기능이 추가되었습니다. 자세한 내용은 Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용 및 Amazon EMR에서 Spark와 함께 AWS Glue 데이터 카탈로그 사용 섹션을 참조하세요.
- YARN 애플리케이션에 대한 기록 데이터 및 Spark 애플리케이션에 대한 추가 세부 정보를 볼 수 있는 클러스터 정보에 애플리케이션 이력을 추가했습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 애플리케이션 기록 보기를 참조하세요.
- Oozie
  - OOZIE-2748을 백포트했습니다.
- Hue
  - HUE-5859를 백포트했습니다
- HBase
  - `getMasterInitializedTime`을 사용하여 JMX(Java Management Extensions)를 통해 HBase 마스터 서버 시작 시간을 공개하도록 패치를 추가했습니다.
  - 클러스터 시작 시간을 개선하는 패치를 추가했습니다.

알려진 문제

- 모든 애플리케이션을 설치하고 기본 Amazon EBS 루트 볼륨 크기를 변경하지 않은 상태에서 클러스터가 시작되지 않습니다. 해결 방법으로의 `aws emr create-cluster` 명령을 사용하고 더 큰 `--ebs-root-volume-size` 파라미터를 AWS CLI 지정합니다.
- Hive 2.3.0에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 [Statistics in Hive]를 참조하십시오.
- Spark - Spark를 사용할 때, apppusher 대몬(daemon)과 관련된 파일 핸들러 유출 문제가 있습니다. 이는 몇 시간 또는 며칠 후에 장시간 실행되는 Spark 작업에 나타날 수 있습니다. 이 문제를 해결하려면 프라이머리 노드에 연결하고 `sudo /etc/init.d/apppusher stop`을 입력합니다. 이렇게 하면 apppusher 대몬(daemon)이 중단되며 Amazon EMR이 이를 자동으로 다시 시작합니다.
- 애플리케이션 이력
  - 작동 중단 Spark 실행기의 기록 데이터는 사용할 수 없습니다.
  - 진행 중인 암호화를 사용하기 위해 보안 구성을 사용하는 클러스터에서는 애플리케이션 이력을 사용할 수 없습니다.

5.8.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*-amzn-*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.4.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.4.0	편리한 하둡 에코시스템용 추가 라이브러리.
emr-kinesis	3.4.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.6.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.18.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.7.3-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.

구성 요소	버전	설명
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.3.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.3.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.3.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.3.0-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.3.0-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.3.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.3.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.

구성 요소	버전	설명
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-1	Pig 명령줄 클라이언트.
spark-client	2.2.0	Spark 명령줄 클라이언트.
spark-history-server	2.2.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.2.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.2.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.8.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-5.8.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.7.1

5.7.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Mahout, HueOozie, PhoenixPig, Presto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.0	1.3.0	1.2.1	1.2.1
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.1	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.13.0
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.11.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-5.7.1	emr-5.7.0	emr-5.6.1	emr-5.6.0
Zeppelin	0.7.2	0.7.2	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.7.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.7.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리

구성 요소	버전	설명
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.18.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.

구성 요소	버전	설명
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.1	Spark 명령줄 클라이언트.
spark-history-server	2.1.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.7.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.7.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.7.0

5.7.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Mahout, HueOozie, PhoenixPig, Presto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.3.0	1.2.1	1.2.1	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.1	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.7.0	emr-5.6.1	emr-5.6.0	emr-5.5.4
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.13.0	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.11.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.1	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	0.7.2	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.7.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.7.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.6.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 7월 13일

업그레이드

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

새로운 특성

- 클러스터를 생성할 때 사용자 지정 Amazon Linux AMI를 지정하는 기능을 추가했습니다. 자세한 내용은 사용자 지정 AMI 사용을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- HBase
  - HBase 읽기 전용 복제본 클러스터를 구성하는 기능을 추가했습니다. 읽기 전용 복제본 클러스터 사용을 참조하십시오.
  - 다양한 버그 수정 사항 및 기능 향상
- Presto - `node.properties` 구성 기능을 추가했습니다.
- YARN - `container-log4j.properties` 구성 기능을 추가했습니다.
- Sqoop - SQOOP-2880을 백포트하여 Sqoop 임시 디렉터리를 설정할 수 있는 인수를 도입했습니다.

5.7.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.18.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.3.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.

구성 요소	버전	설명
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.

구성 요소	버전	설명
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.11.0-HBase-1.3	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.11.0-HBase-1.3	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.1	Spark 명령줄 클라이언트.
spark-history-server	2.1.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.7.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.7.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.7.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.

분류	설명
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.6.1

5.6.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Mahout, HueOozie, PhoenixPig, Presto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.1	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.13.0	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.1	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-

	emr-5.6.1	emr-5.6.0	emr-5.5.4	emr-5.5.3
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.6.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.6.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리

구성 요소	버전	설명
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.17.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.2.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.

구성 요소	버전	설명
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.1	Spark 명령줄 클라이언트.
spark-history-server	2.1.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.1	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.6.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.6.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.6.0

5.6.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Mahout, HueOozie, PhoenixPig, Presto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.1	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.6.0	emr-5.5.4	emr-5.5.3	emr-5.5.2
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.13.0	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.1	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.6.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.6.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.5.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 6월 5일

업그레이드

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Amazon EMR 버전 5.0 이상에서 Spark 2.x를 지원하는 첫 번째 버전의 Mahout입니다.
- Spark 2.1.1

변경 사항, 향상된 기능 및 해결된 문제

- Presto
  - 보안 구성에서 전송 데이터 암호화를 사용하여 Presto 노드 간 SSL/TLS 보안 통신을 활성화할 수 있는 기능이 추가되었습니다. 자세한 내용은 [전송 중인 데이터 암호화](#)를 참조하십시오.
  - [Presto 7661](#)의 백포트와 함께 `VERBOSE` 옵션이 `EXPLAIN ANALYZE` 문에 추가되면서 쿼리 계획에 대해 더욱 자세한 저레벨 통계를 보고할 수 있게 되었습니다.

5.6.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터

구성 요소	버전	설명
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.17.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.2.1	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리

구성 요소	버전	설명
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스.
hbase-region-server	1.3.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스.
hbase-client	1.3.1	HBase 명령줄 클라이언트.
hbase-rest-server	1.3.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스.

구성 요소	버전	설명
hbase-thrift-server	1.3.1	HBase에 Thrift 엔드포인트를 제공하는 서비스.
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.13.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.

구성 요소	버전	설명
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.1	Spark 명령줄 클라이언트.
spark-history-server	2.1.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.1	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.6.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.6.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-env	Presto의 presto-env.sh 파일에서 값을 변경합니다.
presto-node	Presto의 node.properties 파일에서 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.5.4

5.5.4 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-

	emr-5.5.4	emr-5.5.3	emr-5.5.2	emr-5.5.1
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.5.4 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 [새 소식]을 참조하십시오.

5.5.4 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.16.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.2.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.3.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.3.0	HBase 명령줄 클라이언트
hbase-rest-server	1.3.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.3.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.5.4 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-5.5.4 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.5.3

5.5.3 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPhoenixPig, Presto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.12.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.5.3	emr-5.5.2	emr-5.5.1	emr-5.5.0
Presto	0.170	0.170	0.170	0.170
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.1
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.10

5.5.3 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.3에 대한 정보가 포함됩니다. 변경 사항은 5.5.2에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

5.5.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.16.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.2.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.3.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	1.3.0	HBase 명령줄 클라이언트
hbase-rest-server	1.3.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.3.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.

구성 요소	버전	설명
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.5.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.5.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.5.2

5.5.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.12.0	3.11.0
Iceberg	-	-	-	-

	emr-5.5.2	emr-5.5.1	emr-5.5.0	emr-5.4.1
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.170	0.170	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.1	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.10	3.4.9

5.5.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.2에 대한 정보가 포함됩니다. 변경 사항은 5.5.1에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

5.5.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.16.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.

구성 요소	버전	설명
flink-client	1.2.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.3.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.3.0	HBase 명령줄 클라이언트
hbase-rest-server	1.3.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.3.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.5.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.5.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.5.1

5.5.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.2.0
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.3.0
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.12.0	3.12.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.9.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.5.1	emr-5.5.0	emr-5.4.1	emr-5.4.0
Presto	0.170	0.170	0.166	0.166
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.1	0.7.0	0.7.0
ZooKeeper	3.4.10	3.4.10	3.4.9	3.4.9

5.5.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.5.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

5.5.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.16.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.2.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.3.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.3.0	HBase 명령줄 클라이언트
hbase-rest-server	1.3.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.3.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.

구성 요소	버전	설명
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.

구성 요소	버전	설명
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.5.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.5.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.

분류	설명
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.

분류	설명
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.

분류	설명
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.5.0

5.5.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.2.0	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.3.0	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-

	emr-5.5.0	emr-5.4.1	emr-5.4.0	emr-5.3.2
Hue	3.12.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.9.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.170	0.166	0.166	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.1	0.7.0	0.7.0	0.6.2
ZooKeeper	3.4.10	3.4.9	3.4.9	3.4.9

5.5.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.4.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 4월 26일

업그레이드

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1
- ZooKeeper 3.4.10

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - 외부 셔플 서비스를 사용할 수 없을 때 손실된 셔플 블록을 모두 재계산하기 위한 Spark Patch(SPARK-20115) Fix DAGScheduler가 이번 릴리스에 포함되어 있는 Spark 2.1.0으로 백포트되었습니다.
- Flink
  - Flink는 이제 Scala 2.11로 빌드됩니다. 따라서 Scala API 및 라이브러리를 사용하는 경우에는 프로젝트에 Scala 2.11을 사용하는 것이 좋습니다.
  - `HADOOP_CONF_DIR` 및 `YARN_CONF_DIR` 기본값이 올바로 설정되지 않아서 `start-scala-shell.sh` 작업이 중단되던 문제가 해결되었습니다. 또한 `env.hadoop.conf.dir` 또는 `env.yarn.conf.dir` 구성 분류에서 `/etc/flink/conf/flink-conf.yaml` 및 `flink-conf`을 사용하여 이 값들을 설정할 수 있는 기능도 추가되었습니다.
  - 새로운 EMR 전용 명령인 `flink-scala-shell`이 `start-scala-shell.sh` 래퍼로 도입되었습니다. 앞으로 `start-scala-shell`이 아닌 이 명령을 사용하는 것이 좋습니다. 새로운 명령은 실행이 간편합니다. 예를 들어 `flink-scala-shell -n 2`는 Flink Scala 셸을 작업 병렬 처리 2로 시작합니다.
  - 새로운 EMR 전용 명령인 `flink-yarn-session`이 `yarn-session.sh` 래퍼로 도입되었습니다. 앞으로 `yarn-session`이 아닌 이 명령을 사용하는 것이 좋습니다. 새로운 명령은 실행이 간편합니다. 예를 들어, `flink-yarn-session -d -n 2`는 연결이 끊어진 상태에서 태스크 관리자 2개를 사용하여 장기 실행 Flink 세션을 시작합니다.
  - (FLINK-6125) 공통 httpclient가 더 이상 Flink 1.2에서 셰이딩되지 않는 문제가 해결되었습니다.

- Presto
  - LDAP 인증에 대한 지원이 추가되었습니다. 단, Amazon EMR 기반 Presto에서 LDAP를 사용하려면 Presto 코디네이터의 HTTPS 액세스를 활성화해야 합니다(`config.properties`의 `http-server.https.enabled=true`). 구성에 대한 자세한 내용은 Presto 설명서에서 LDAP Authentication을 참조하십시오.
  - `SHOW GRANTS` 지원이 추가되었습니다.
- Amazon EMR 기본 Linux AMI
  - Amazon EMR 릴리스는 이제 Amazon Linux 2017.03을 기반으로 합니다. 자세한 내용은 Amazon Linux AMI 2017.03 릴리스 정보를 참조하십시오.
  - Amazon EMR 기본 Linux 이미지에서 Python 2.6을 제거했습니다. 이제는 Python 2.7 및 3.4가 기본 설치됩니다. 필요하다면 Python 2.6을 수동으로 설치할 수도 있습니다.

5.5.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.5.0	Amazon S3용으로 최적화된 분산 복사 애플리케이션.
emrfs	2.16.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.2.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.3.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.3.0	HBase 명령줄 클라이언트
hbase-rest-server	1.3.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.3.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.12.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.170	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.170	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.1	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.10	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.10	ZooKeeper 명령줄 클라이언트.

5.5.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.5.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.

분류	설명
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.

분류	설명
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.4.1

5.4.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.2.0	1.2.0	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.3.0	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.4.1	emr-5.4.0	emr-5.3.2	emr-5.3.1
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.9.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.166	0.166	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.7.0	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

5.4.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.4.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.15.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.2.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.3.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.3.0	HBase 명령줄 클라이언트
hbase-rest-server	1.3.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.3.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스

구성 요소	버전	설명
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.166	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.166	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.

구성 요소	버전	설명
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.4.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-5.4.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.

분류	설명
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.

분류	설명
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.

분류	설명
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.

분류	설명
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.

분류	설명
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.4.0

5.4.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
Flink	1.2.0	1.1.4	1.1.4	1.1.4
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.3.0	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.1
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.1
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.11.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.3.0
Phoenix	4.9.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.166	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6

	emr-5.4.0	emr-5.3.2	emr-5.3.1	emr-5.3.0
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.7.0	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

5.4.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.4.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.3.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 3월 8일

업그레이드

- Flink 1.2.0로 업그레이드
- HBase 1.3.0으로 업그레이드
- Phoenix 4.9.0로 업그레이드

> **Note**
> 이전 버전의 Amazon EMR에서 Amazon EMR 버전 5.4.0 이상으로 업그레이드하고 보조 인덱싱을 사용하는 경우 Apache Phoenix 설명서에 설명된 대로 로컬 인덱스를 업그레이드합니다. Amazon EMR은 `hbase-site` 분류에서 필수 구성을 제거하지만 인덱스를 다시 채워야 합니다. 인덱스의 온라인 및 오프라인 업그레이드가 지원됩니다. 온라인 업그레이드는 기본값이므로, 버전 4.8.0 이상인 Phoenix 클라이언트에서 초기화하는 동안 인덱스가 다시 채워집니다. 오프라인 업그레이드를 지정하려면 `phoenix.client.localIndexUpgrade` 분류에서 `phoenix-site` 구성을 false로 설정한 다음, SSH를 프라이머리 노드로 설정하여 `psql [zookeeper] -1`을 실행합니다.

- Presto 0.166로 업그레이드

- Zeppelin 0.7.0로 업그레이드

변경 사항 및 기능 향상

- r4 인스턴스에 대한 지원을 추가했습니다. Amazon EC2 인스턴스 유형을 참조하세요.

5.4.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.3.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.15.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.

구성 요소	버전	설명
flink-client	1.2.0	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.3.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.3.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.3.0	HBase 명령줄 클라이언트
hbase-rest-server	1.3.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.3.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-hbase	2.1.1-amzn-0	Hive-hbase 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server2	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.9.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.9.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.166	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.166	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.7.0	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.4.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.4.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.3.2

5.3.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.4	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.1	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.1	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.11.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.3.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.3.2	emr-5.3.1	emr-5.3.0	emr-5.2.3
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.1.0	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

5.3.2 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.3.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.14.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.4	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.

구성 요소	버전	설명
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.3.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.3.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.3.1

5.3.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.4	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.1	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.1	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.11.0	3.10.0	3.10.0
Iceberg	-	-	-	-

	emr-5.3.1	emr-5.3.0	emr-5.2.3	emr-5.2.2
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.3.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.1.0	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

5.3.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.3.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.3.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 2월 7일

Zeppelin 패치를 백포트하고 Amazon EMR의 기본 AMI를 업데이트하는 부수적 변경 내용을 포함합니다.

5.3.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.14.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.4	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.1.1-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.11.0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.3.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.3.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.3.0

5.3.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.4	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.1	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.1	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.11.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.3.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.3.0	emr-5.2.3	emr-5.2.2	emr-5.2.1
Presto	0.157.1	0.157.1	0.157.1	0.157.1
Spark	2.1.0	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.9

5.3.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.3.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.2.1 릴리스에 관련됩니다.

릴리스 날짜: 2017년 1월 26일

업그레이드

- Hive 2.1.1으로 업그레이드
- Hue 3.11.0로 업그레이드
- Spark 2.1.0로 업그레이드
- Oozie 4.3.0로 업그레이드
- Flink 1.1.4로 업그레이드

변경 사항 및 기능 향상

- `interpreters_shown_on_wheel` 파일에 지정된 순서와 상관없이 `hue.ini` 설정을 사용하여 노트북 선택 휠에 첫 번째로 표시할 인터프리터를 구성할 수 있는 Hue 패치를 추가했습니다.

- Hive의 `hive-parquet-logging` 파일에서 값을 구성하는 데 사용할 수 있는 `parquet-logging.properties` 구성 분류를 추가했습니다.

5.3.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.14.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.4	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.1-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.1-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.1-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.1-amzn-0	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.1.1-amzn-0	하듑 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.1-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.11.0	하듑 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.3.0	Oozie 명령줄 클라이언트.
oozie-server	4.3.0	Oozie 워크플로 요청을 수락하는 서비스.
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.1.0	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.1.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.1.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.1.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.3.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.3.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-parquet-logging	Hive parquet-logging.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.2.3

5.2.3 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.2.3	emr-5.2.2	emr-5.2.1	emr-5.2.0
Presto	0.157.1	0.157.1	0.157.1	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.9	3.4.8

5.2.3 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.2.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.13.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.3	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트

구성 요소	버전	설명
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.2	Spark 명령줄 클라이언트.
spark-history-server	2.0.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.2.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-5.2.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.

분류	설명
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.

분류	설명
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.

분류	설명
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.2.2

5.2.2 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전

- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-

	emr-5.2.2	emr-5.2.1	emr-5.2.0	emr-5.1.1
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.157.1	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.2	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.9	3.4.8	3.4.8

5.2.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.2.2에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.2.1 릴리스에 관련됩니다.

릴리스 날짜: 2017년 5월 2일

이전 릴리스에서 해결된 알려진 문제

- SPARK-194459의 백포트와 함께 char/varchar 열이 포함된 ORC 테이블에서 데이터를 읽어오지 못하던 문제가 해결되었습니다.

5.2.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.13.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.3	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.2	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.0.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.2.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.2.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.2.1

5.2.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	1.1.3
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.3
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.157.1	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.2	2.0.1	2.0.1

	emr-5.2.1	emr-5.2.0	emr-5.1.1	emr-5.1.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.2
ZooKeeper	3.4.9	3.4.8	3.4.8	3.4.8

5.2.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.2.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.2.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 12월 29일

업그레이드

- Presto 0.157.1로 업그레이드되었습니다. 자세한 내용은 Presto 설명서의 Presto 릴리스 정보를 참조하십시오.
- Zookeeper 3.4.9로 업그레이드되었습니다. 자세한 내용은 Apache ZooKeeper 설명서의 ZooKeeper 릴리스 정보를 참조하십시오.

변경 사항 및 기능 향상

- Amazon EMR 버전 4.8.3 이상(5.0.0, 5.0.3 및 5.2.0 제외)에서 Amazon EC2 m4.16xlarge 인스턴스 유형에 대한 지원을 추가했습니다.
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.
- 이제 Flink 및 YARN 구성 경로가 `/etc/default/flink`에서 기본적으로 설정되므로 `FLINK_CONF_DIR` 또는 `HADOOP_CONF_DIR` 드라이버 스크립트를 실행하여 Flink 작업을 시작할 때 환경 변수 `flink` 및 `yarn-session.sh`을 설정할 필요가 없습니다.

- FlinkKinesisConsumer 클래스에 대한 지원을 추가했습니다.

이전 릴리스에서 해결된 알려진 문제

- 대규모 클러스터에서 동일한 파일의 복제 및 삭제 간 경합 때문에 ReplicationMonitor 스레드가 장시간 정체될 수 있는 하둡 문제를 수정했습니다.
- 작업 상태가 성공적으로 업데이트되지 않은 경우 ControlledJob#toString이 null 포인터 예외(NPE)와 함께 실패하는 문제를 수정했습니다.

5.2.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.13.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.3	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.2	Spark 명령줄 클라이언트.
spark-history-server	2.0.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

5.2.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.2.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.2.0

5.2.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	1.1.3	1.1.3	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.3	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.2.0	emr-5.1.1	emr-5.1.0	emr-5.0.3
Presto	0.152.3	0.152.3	0.152.3	0.152.3
Spark	2.0.2	2.0.1	2.0.1	2.0.1
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.2	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

5.2.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.2.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.1.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 11월 21일

변경 사항 및 기능 향상

- HBase에 대한 Amazon S3 스토리지 모드를 추가했습니다.
- HBase rootdir의 Amazon S3 위치를 지정할 수 있습니다. 자세한 내용은 Amazon S3용 HBase를 참조하세요.

업그레이드

- Spark 2.0.2로 업그레이드

이전 릴리스에서 해결된 알려진 문제

- EBS 전용 인스턴스 유형에서 /mnt가 2TB로 제약되는 문제를 수정했습니다.

- instance-controller 및 logpusher 로그가 매시간 회전하는 정상적인 og4j-configured .log 파일 대신 해당 .out 파일로 출력되는 문제를 수정했습니다. .out 파일은 회전하지 않으므로 결국 /emr 파티션이 파일로 가득 찰 수 있습니다. 이 문제는 하드웨어 가상 머신(HVM) 인스턴스 유형에만 영향을 미칩니다.

5.2.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.12.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.3	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.7.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.152.3	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.152.3	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.2	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	2.0.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

5.2.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.2.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase	Apache HBase에 대한 Amazon EMR 큐레이트 설정입니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.1.1

5.1.1 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
Delta	-	-	-	-
Flink	1.1.3	1.1.3	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.3	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.152.3	0.150
Spark	2.0.1	2.0.1	2.0.1	2.0.0

	emr-5.1.1	emr-5.1.0	emr-5.0.3	emr-5.0.2
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.2	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

5.1.1 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

5.1.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.11.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.3	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.7.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.3	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스

구성 요소	버전	설명
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.152.3	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.152.3	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.1	Spark 명령줄 클라이언트.
spark-history-server	2.0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

5.1.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.1.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.

분류	설명
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.1.0

5.1.0 애플리케이션 버전

이 릴리스에는 Flink, , Ganglia, HBase, HCatalog, Hadoop, , Hive, Hue, Oozie, MahoutPig, PhoenixPresto, Spark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	1.1.3	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.3	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-5.1.0	emr-5.0.3	emr-5.0.2	emr-5.0.1
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.152.3	0.150	0.150
Spark	2.0.1	2.0.1	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.2	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

5.1.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.1.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.0.3 릴리스에 관련됩니다.

릴리스 날짜: 2016년 11월 3일

변경 사항 및 기능 향상

- Flink 1.1.3에 대한 지원을 추가했습니다.
- Presto는 Hue의 노트북 섹션에서 옵션으로 추가되었습니다.

업그레이드

- HBase 1.2.3로 업그레이드
- Zeppelin 0.6.2로 업그레이드

이전 릴리스에서 해결된 알려진 문제

- ORC 파일이 이전 Amazon EMR 4.x 버전만큼 좋은 성능을 나타내지 않는 Amazon S3의 Tez 쿼리 문제를 수정했습니다.

5.1.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.11.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
flink-client	1.1.3	Apache Flink 명령줄 클라이언트 스크립트 및 애플리케이션
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.7.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.3	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.3	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.3	HBase 명령줄 클라이언트
hbase-rest-server	1.2.3	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.3	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.152.3	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
presto-worker	0.152.3	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.1	Spark 명령줄 클라이언트.
spark-history-server	2.0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.2	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

5.1.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.1.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
flink-conf	flink-conf.yaml 설정을 변경합니다.
flink-log4j	Flink log4j.properties 설정을 변경합니다.
flink-log4j-yarn-session	Flink log4j-yarn-session.properties 설정을 변경합니다.
flink-log4j-cli	Flink log4j-cli.properties 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.

분류	설명
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.0.3

5.0.3 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, OoziePhoenix, Presto, PigSpark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

5.0.3 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 5.0.3에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.0.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 10월 24일

업그레이드

- Hadoop 2.7.3로 업그레이드
- Presto 0.152.3로 업그레이드했습니다. 여기에는 Presto 웹 인터페이스에 대한 지원이 포함됩니다. 포트 8889를 사용하여 Presto 코디네이터의 Presto 웹 인터페이스에 액세스할 수 있습니다. Presto 웹 인터페이스에 대한 자세한 내용은 Presto 설명서의 웹 인터페이스를 참조하십시오.
- Spark 2.0.1으로 업그레이드
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.

5.0.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.10.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.7.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.

구성 요소	버전	설명
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.152.3	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.152.3	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.1	Spark 명령줄 클라이언트.
spark-history-server	2.0.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

5.0.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.0.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.

분류	설명
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.

분류	설명
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.0.2

5.0.2 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, OoziePhoenix, Presto, PigSpark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Scala	2.11.8	2.11.8	2.11.8	2.11.8
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

5.0.2 릴리스 정보

5.0.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.9.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리

구성 요소	버전	설명
hadoop-hdfs-namenode	2.7.2-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.150	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.150	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.0	Spark 명령줄 클라이언트.
spark-history-server	2.0.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.1-SNAPSHOT	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

5.0.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.0.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.0.1

5.0.1 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, OoziePhoenix, Presto, PigSpark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

5.0.1 릴리스 정보

5.0.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터

구성 요소	버전	설명
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.9.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스

구성 요소	버전	설명
hadoop-httpfs-server	2.7.2-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.150	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.150	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.0	Spark 명령줄 클라이언트.
spark-history-server	2.0.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.1-SNAPSHOT	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

5.0.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.0.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.

분류	설명
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 5.0.0

5.0.0 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, OoziePhoenix, Presto, PigSpark, Sqoop, 및 애플리케이션이 포함되어 Tez Zeppelin있습니다ZooKeeper.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	2.11.8	2.11.8	2.11.8	2.11.8

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	2.1.0	2.1.0	2.1.0	2.1.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	2.1.0	2.1.0	2.1.0	2.1.0
Hudi	-	-	-	-
Hue	3.10.0	3.10.0	3.10.0	3.10.0
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.16.0	0.16.0	0.16.0	0.16.0

	emr-5.0.3	emr-5.0.2	emr-5.0.1	emr-5.0.0
Presto	0.152.3	0.150	0.150	0.150
Spark	2.0.1	2.0.0	2.0.0	2.0.0
Sqoop	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	3.4.8	3.4.8	3.4.8	3.4.8

5.0.0 릴리스 정보

릴리스 날짜: 2016년 7월 27일

업그레이드

- Hive 2.1으로 업그레이드
- Presto 0.150로 업그레이드
- Spark 2.0로 업그레이드
- Hue 3.10.0로 업그레이드
- Pig 0.16.0로 업그레이드
- Tez 0.8.4로 업그레이드
- Zeppelin 0.6.1으로 업그레이드

변경 사항 및 기능 향상

- Amazon EMR은 최신 오픈 소스 버전의 Hive(버전 2.1)와 Pig(버전 0.16.0)를 지원합니다. 이전에 Amazon EMR에서 Hive나 Pig를 사용한 적이 있는 경우 일부 사용 사례에 영향을 미칠 수 있습니다. 자세한 내용은 Hive 및 Pig를 참조하십시오.

- Hive 및 Pig용 기본 실행 엔진은 이제 Tez입니다. 이 항목을 변경하려면 각각 `hive-site` 및 `pig-properties` 구성 분류에서 해당 값을 편집합니다.
- 향상된 단계 디버깅 기능이 추가되었습니다. 서비스에서 원인을 결정할 수 있는 경우 이 기능을 사용하여 단계 실패의 근본 원인을 확인할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 향상된 단계 디버깅을 참조하세요.
- 이전에 "-Sandbox"로 끝났던 애플리케이션은 더 이상 해당 접미사를 사용하지 않습니다. 이로 인해 예를 들어 스크립트를 사용하여 이 애플리케이션이 있는 클러스터를 시작하는 경우 자동화가 중단될 수 있습니다. 다음 테이블에서는 Amazon EMR 5.0.0 대비 Amazon EMR 4.7.2의 애플리케이션 이름을 보여줍니다.

  애플리케이션 이름 변경

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin
ZooKeeper-Sandbox	ZooKeeper

- Spark는 이제 Scala 2.11용으로 컴파일됩니다.
- 이제 기본 JVM은 Java 8입니다. 모든 애플리케이션은 Java 8 실행 시간을 사용하여 실행됩니다. 애플리케이션의 바이트 코드 대상에는 변경 사항이 없습니다. 대부분의 애플리케이션은 계속해서 Java 7을 대상으로 합니다.
- Zeppelin에는 이제 인증 기능이 포함됩니다. 자세한 내용은 Zeppelin을 참조하십시오.
- 암호화 옵션을 더 쉽게 생성하고 적용할 수 있도록 보안 구성에 대한 지원을 추가했습니다. 자세한 내용은 데이터 암호화를 참조하십시오.

5.0.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.9.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.7.2-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	2.1.0-amzn-0	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	2.1.0-amzn-0	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	2.1.0-amzn-0	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	2.1.0-amzn-0	Hive 명령줄 클라이언트
hive-metastore-server	2.1.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	2.1.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.10.0-amzn-0	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.

구성 요소	버전	설명
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.150	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.150	여러 쿼리를 실행하는 서비스.
pig-client	0.16.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	2.0.0	Spark 명령줄 클라이언트.
spark-history-server	2.0.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	2.0.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	2.0.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.1-SNAPSHOT	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

5.0.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-5.0.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.

분류	설명
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j2	HCatalog WebHCat log4j2.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-beeline-log4j2	Hive beeline-log4j2.properties 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.

분류	설명
hive-exec-log4j2	Hive hive-exec-log4j2.properties 파일의 값을 변경합니다.
hive-llap-daemon-log4j2	Hive llap-daemon-log4j2.properties 파일의 값을 변경합니다.
hive-log4j2	Hive hive-log4j2.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-hive-site	Spark hive-site.xml 파일의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.

분류	설명
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 4.x 릴리스 버전

이 섹션에는 각 Amazon EMR 4.x 릴리스 버전에 제공되는 애플리케이션 버전, 릴리스 정보, 구성 요소 버전 및 구성 분류가 포함되어 있습니다.

클러스터를 시작할 때 Amazon EMR의 여러 릴리스 중에서 선택할 수 있습니다. 이를 통해 호환성 요구 사항에 맞는 애플리케이션 버전을 테스트하고 사용할 수 있습니다. 릴리스 레이블과 함께 릴리스 번호를 지정합니다. 릴리스 레이블은 `emr-x.x.x` 형식입니다. 예를 들어 `emr-7.9.0`입니다.

새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

모든 Amazon EMR 4.x 릴리스에서 애플리케이션 버전의 전체 테이블은 Amazon EMR 4.x 릴리스의 애플리케이션 버전 섹션을 참조하세요.

주제

- Amazon EMR 4.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스 버전의 차이
- Amazon EMR 릴리스 4.9.6
- Amazon EMR 릴리스 4.9.5
- Amazon EMR 릴리스 4.9.4
- Amazon EMR 릴리스 4.9.3
- Amazon EMR 릴리스 4.9.2

- Amazon EMR 릴리스 4.9.1
- Amazon EMR 릴리스 4.8.5
- Amazon EMR 릴리스 4.8.4
- Amazon EMR 릴리스 4.8.3
- Amazon EMR 릴리스 4.8.2
- Amazon EMR 릴리스 4.8.1
- Amazon EMR 릴리스 4.8.0
- Amazon EMR 릴리스 4.7.4
- Amazon EMR 릴리스 4.7.3
- Amazon EMR 릴리스 4.7.2
- Amazon EMR 릴리스 4.7.1
- Amazon EMR 릴리스 4.7.0
- Amazon EMR 릴리스 4.6.1
- Amazon EMR 릴리스 4.6.0
- Amazon EMR 릴리스 4.5.0
- Amazon EMR 릴리스 4.4.0
- Amazon EMR 릴리스 4.3.0
- Amazon EMR 릴리스 4.2.0
- Amazon EMR 릴리스 4.1.0
- Amazon EMR 릴리스 4.0.0

Amazon EMR 4.x 릴리스의 애플리케이션 버전

각 Amazon EMR 4.x 릴리스에서 사용할 수 있는 애플리케이션 버전을 표시한 전체 테이블을 보려면 브라우저에서 Amazon EMR 4.x 릴리스의 애플리케이션 버전을 엽니다.

Amazon EMR 4.x 릴리스 버전의 차이

Amazon EMR 관리 안내서의 Amazon EMR 기능 설명서에는 기능을 처음 제공한 Amazon EMR 릴리스 버전과 4.0.0까지 Amazon EMR 기능 간 해당되는 차이점이 나와 있습니다.

Amazon EMR 릴리스 버전 5.0.0부터 일부 애플리케이션에 설치 또는 작동 세부 정보를 변경하는 중대한 버전 업그레이드가 있었으며 샌드박스 애플리케이션에서 네이티브 애플리케이션으로 승격된 애플

리케이션도 있습니다. 이 섹션의 각 주제에서는 Amazon EMR 4.x 릴리스 버전을 사용할 때 나타나는 주요 애플리케이션별 차이를 설명합니다.

주제

샌드박스 애플리케이션

Amazon EMR 4.x 릴리스 버전을 사용할 때 일부 애플리케이션은 샌드박스 애플리케이션으로 간주됩니다. 샌드박스 애플리케이션은 필요에 의해 처음 Amazon EMR 출시 당시 제공된 애플리케이션의 초기 버전입니다. 콘솔 AWS CLI또는 API를 사용하여 Amazon EMR이 네이티브 애플리케이션과 동일한 방식으로 샌드박스 애플리케이션을 설치하도록 할 수 있지만 샌드박스 애플리케이션은 지원 및 설명서가 제한적입니다. Amazon EMR 릴리스 버전 5.0.0 이상에서 샌드박스 애플리케이션은 완벽하게 지원되는 네이티브 애플리케이션이 되었습니다. 다음은 Amazon EMR 4.x 릴리스 버전의 샌드박스 애플리케이션입니다.

- Oozie
- Presto
- Sqoop
- Zeppelin
- ZooKeeper

샌드박스 애플리케이션을 설치할 때 애플리케이션 이름에 접미사 `-sandbox`가 붙습니다. 예를 들어, *`Presto`*의 샌드박스 버전을 설치하려면 `Presto-sandbox`를 사용하십시오. 완벽하게 지원되는 애플리케이션에 비해 설치 시간이 오래 걸릴 수 있습니다. 이 단원에 나오는 각 애플리케이션의 버전 번호는 애플리케이션의 커뮤니티 버전에 해당합니다.

Oozie(샌드박스 버전)

Amazon EMR 릴리스 버전 4.1.0부터 Oozie가 샌드박스 애플리케이션으로 제공됩니다.

샌드박스 버전을 사용하면 Oozie 예제가 기본적으로 설치되지 않습니다. 예제를 설치하려면 마스터 노드에 SSH하고 `install-oozie-examples`를 실행합니다.

Oozie-Sandbox 버전 정보

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.9.6	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
		cemanager, oozie-client, oozie-server
emr-4.9.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.9.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.9.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.5	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.8.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.8.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.8.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.8.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.4	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.7.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.7.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.7.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.6.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.6.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.5.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.4.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Oozie-Sandbox 버전	Oozie-Sandbox와 함께 설치된 구성 요소
emr-4.3.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server
emr-4.1.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, oozie-client, oozie-server

Presto(샌드박스 버전)

Amazon EMR 릴리스 버전 4.1.0부터 Presto가 샌드박스 애플리케이션으로 제공됩니다.

Presto-Sandbox 버전 정보

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.9.6	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server,

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
		presto-coordinator, presto-worker
emr-4.9.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.9.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.9.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.5	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.4	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.8.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.2	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-4.8.1	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.8.0	0.151	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-4.7.4	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-4.7.3	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.7.2	0.148	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-4.7.1	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-4.7.0	0.147	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.6.1	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.6.0	0.143	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.5.0	0.140	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.4.0	0.136	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.3.0	0.130	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker
emr-4.2.0	0.125	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto-Sandbox 버전	Presto-Sandbox와 함께 설치된 구성 요소
emr-4.1.0	0.119	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, mysql-server, presto-coordinator, presto-worker

Sqoop(샌드박스 버전)

Amazon EMR 릴리스 버전 4.4.0부터 Sqoop이 샌드박스 애플리케이션으로 제공됩니다.

Sqoop-Sandbox 버전 정보

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
emr-4.9.6	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
		httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
emr-4.9.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.5	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
emr-4.8.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
emr-4.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
emr-4.7.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
emr-4.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-4.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop-Sandbox 버전	Sqoop-Sandbox와 함께 설치된 구성 요소
emr-4.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client
emr-4.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, sqoop-client

Zeppelin(샌드박스 버전)

Amazon EMR 릴리스 버전 4.1.0부터 Zeppelin이 샌드박스 애플리케이션으로 제공됩니다.

Zeppelin-Sandbox 버전 정보

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.9.6	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
		nager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.9.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.9.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.9.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.5	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.8.4	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.8.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.8.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.8.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.4	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.7.3	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.2	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.7.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.7.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.6.1	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.6.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.5.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.4.0	0.5.6	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin-Sandbox 버전	Zeppelin-Sandbox와 함께 설치된 구성 요소
emr-4.3.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.2.0	0.5.5	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-4.1.0	0.6.0-SNAPSHOT	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Zookeeper(샌드박스 버전)

Amazon EMR 릴리스 버전 4.6.0부터 Zookeeper가 샌드박스 애플리케이션으로 제공됩니다.

ZooKeeper-Sandbox 버전 정보

Amazon EMR 릴리스 레이블	ZooKeeper-Sandbox 버전	ZooKeeper-Sandbox와 함께 설치된 구성 요소
emr-4.9.6	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper-Sandbox 버전	ZooKeeper-Sandbox와 함께 설치된 구성 요소
emr-4.9.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.9.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper-Sandbox 버전	ZooKeeper-Sandbox와 함께 설치된 구성 요소
emr-4.8.5	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.4	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.8.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper-Sandbox 버전	ZooKeeper-Sandbox와 함께 설치된 구성 요소
emr-4.8.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, zookeeper-client, zookeeper-server
emr-4.8.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, zookeeper-client, zookeeper-server
emr-4.8.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper-Sandbox 버전	ZooKeeper-Sandbox와 함께 설치된 구성 요소
emr-4.7.4	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper-Sandbox 버전	ZooKeeper-Sandbox와 함께 설치된 구성 요소
emr-4.7.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.7.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-4.6.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper-Sandbox 버전	ZooKeeper-Sandbox와 함께 설치된 구성 요소
emr-4.6.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 4.x에서 Hive 사용 시 고려 사항

이 섹션에서는 Amazon EMR 5.x 릴리스 버전에서 Hive 2.x를 사용하는 것에 비해 Amazon EMR 4.x 릴리스 버전에서 Hive 버전 1.0.0을 사용할 때 고려할 차이점을 설명합니다.

ACID 트랜잭션 지원되지 않음

Amazon EMR 4.x 릴리스 버전의 Hive는 4.x 릴리스 버전을 사용할 때 Hive 데이터가 Amazon S3에 저장되어 ACID 트랜잭션이 지원되지 않습니다. Amazon S3에서 트랜잭션 테이블을 생성하려고 하면 예외가 발생합니다.

Amazon S3에서 테이블에 대한 읽기 및 쓰기

Amazon EMR 4.x 릴리스 버전의 Hive는 임시 파일을 사용하지 않고 Amazon S3에 직접 쓸 수 있습니다. 이 기능으로 인해 성능이 향상되지만 같은 Hive 문에서 Amazon S3 내 동일한 테이블에 대한 읽기 및 쓰기 작업을 수행할 수 없습니다. HDFS에서 임시 테이블을 생성하여 사용하는 차선책이 있습니다.

다음 예제는 여러 Hive 문을 사용하여 Amazon S3에서 테이블을 업데이트하는 방법을 보여 줍니다. 이 명령문에서는 Amazon S3에 있는 `my_s3_table`이라는 테이블을 기반으로 HDFS에 `tmp`라는 임시 테이블을 생성합니다. 그러면 Amazon S3의 테이블이 임시 테이블 콘텐츠로 업데이트됩니다.

```
CREATE TEMPORARY TABLE tmp LIKE my_s3_table;
INSERT OVERWRITE TABLE tmp SELECT ....;
INSERT OVERWRITE TABLE my_s3_table SELECT * FROM tmp;
```

Log4j 및 Log4j 2 비교

Amazon EMR 4.x 릴리스 버전의 Hive에 Log4j가 사용됩니다. 버전 5.0.0부터는 Log4j 2가 기본값입니다. 이 버전에서는 다른 로깅 구성이 필요할 수 있습니다. 자세한 내용은 Apache Log4j 2를 참조하십시오.

MapReduce가 기본 실행 엔진임

Amazon EMR 4.x 릴리스 버전의 Hive에서는 MapReduce를 기본 실행 엔진으로 사용합니다. Amazon EMR 버전 5.0.0부터는 대다수 워크플로에서 성능을 개선하는 Tez가 기본 실행 엔진입니다.

Hive 권한 부여

Amazon EMR 4.x 릴리스 버전의 Hive는 HDFS에 대해 Hive 권한 부여를 지원하지만 EMRFS 및 Amazon S3에서는 지원하지 않습니다. Amazon EMR 클러스터는 기본적으로 권한 부여가 비활성화된 상태로 실행됩니다.

Amazon S3와 Hive 파일 병합 동작

Amazon EMR 4.x 릴리스 버전의 Hive는 `hive.merge.mapfiles`가 `true`일 경우 맵 전용 작업 끝에서 작은 파일을 병합합니다. 작업의 평균 출력 크기가 `hive.merge.smallfiles.avgsize` 설정보다 작아야 병합이 트리거됩니다. Amazon EMR Hive는 최종 출력 경로가 HDFS에 있는 경우 정확히 동일한 동작을 수행합니다. 하지만 출력 경로가 Amazon S3에 있으면 `hive.merge.smallfiles.avgsize` 파라미터가 무시됩니다. 이 상황에서 `hive.merge.mapfiles`가 `true`로 설정된 경우 병합 작업이 항상 트리거됩니다.

Amazon EMR 4.x에서 Pig 사용 시 고려 사항

Amazon EMR 4.x 릴리스 버전을 사용하여 만든 클러스터에 Pig 버전 0.14.0이 설치됩니다. Pig는 Amazon EMR 5.0.0에서 버전 0.16.0으로 업그레이드되었습니다. 아래에 중대한 차이가 설명되어 있습니다.

다양한 기본 실행 엔진

Amazon EMR 4.x 릴리스 버전의 Pig 버전 0.14.0에서는 MapReduce를 기본 실행 엔진으로 사용합니다. Pig 0.16.0 이상에는 Apache Tez가 사용됩니다. `exectype=mapreduce` 구성 분류에 `pig-properties`를 명시적으로 설정하여 MapReduce를 사용할 수 있습니다.

삭제된 Pig 사용자 정의 함수(UDF)

Amazon EMR 4.x 릴리스 버전의 Pig에 제공된 사용자 지정 UDF가 Pig 0.16.0부터 삭제되었습니다. 대부분의 UDF에는 대신 사용할 수 있는 함수가 있습니다. 다음 표에는 삭제된 UDF와 해당 함수가 있습니다. 자세한 내용은 Apache Pig 사이트에서 Built-in functions를 참조하세요.

삭제된 UDF	동등한 함수
FORMAT_DT(dtformat, 날짜)	GetHour(날짜), GetMinute(날짜), GetMonth(날짜), GetSecond(날짜), GetWeek(날짜), GetYear(날짜), GetDay(날짜)
EXTRACT(문자열, 패턴)	REGEX_EXTRACT_ALL(문자열, 패턴)
REPLACE(문자열, 패턴, 대체)	REPLACE(문자열, 패턴, 대체)
DATE_TIME()	ToDate()
DURATION(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween(dt, dt2), SecondsBetween(dt, dt2), MonthsBetween(dt, dt2), MinutesBetween(dt, dt2), HoursBetween(dt, dt2)
EXTRACT_DT(형식, 날짜)	GetHour(날짜), GetMinute(날짜), GetMonth(날짜), GetSecond(날짜), GetWeek(날짜), GetYear(날짜), GetDay(날짜)
OFFSET_DT(날짜, 기간)	AddDuration(날짜, 기간), SubtractDuration(날짜, 기간)
PERIOD(dt, dt2)	WeeksBetween(dt, dt2), YearsBetween(dt, dt2), SecondsBetween(dt, dt2), MonthsBetween(dt, dt2), MinutesBetween(dt, dt2), HoursBetween(dt, dt2)
CAPITALIZE(문자열)	UCFIRST(문자열)
CONCAT_WITH()	CONCAT()
INDEX_OF()	INDEXOF()

삭제된 UDF	동등한 함수
LAST_INDEX_OF()	LAST_INDEXOF()
SPLIT_ON_REGEX()	STRSPLT()
UNCAPITALIZE()	LCFIRST()

동일한 사항 없는 FORMAT(), LOCAL_DATE(), LOCAL_TIME(), CENTER(), LEFT_PAD(), REPEAT(), REPLACE_ONCE(), RIGHT_PAD(), STRIP(), STRIIP_END(), STRIP_START(), SWAP_CASE()와 같은 UDF는 삭제되었습니다.

사용 중지된 Grunt 명령 사용

Pig 0.16.0부터 일부 Grunt 명령이 중단됩니다. 다음 표에는 Pig 0.14.0의 Grunt 명령과 최신 버전의 동일한 명령(해당되는 경우)이 나와 있습니다.

Pig 0.14.0 명령 및 이와 동등한 최신 Grunt 명령

Pig 0.14.0 Grunt 명령	0.16.0 이상의 Pig Grunt 명령
cat <non-hdfs-path>)	fs -cat <non-hdfs-path>;
cd <non-hdfs-path>;	동일한 사항 없음
ls <non-hdfs-path>;	fs -ls <non-hdfs-path>;
move <non-hdfs-path> <non-hdfs-path>;	fs -mv <non-hdfs-path> <non-hdfs-path>;
copy <non-hdfs-path> <non-hdfs-path>;	fs -cp <non-hdfs-path> <non-hdfs-path>;
copyToLocal <non-hdfs-path> <local-path>;	fs -copyToLocal <non-hdfs-path> <local-path>;
copyFromLocal <local-path> <non-hdfs-path>;	fs -copyFromLocal <local-path> <non-hdfs-path>;
mkdir <non-hdfs-path>;	fs -mkdir <non-hdfs-path>;
rm <non-hdfs-path>;	fs -rm -r -skipTrash <non-hdfs-path>;
rmf <non-hdfs-path>;	fs -rm -r -skipTrash <non-hdfs-path>;

비HDFS 홈 디렉터리에서 제거된 기능

Amazon EMR 4.x 릴리스 버전의 Pig 0.14.0에서는 두 가지 메커니즘을 통해 홈 디렉터리가 없는 hadoop 사용자 이외의 사용자가 Pig 스크립트를 실행할 수 있습니다. 첫 번째 방법은 홈 디렉터리가 존재하지 않는 경우 초기 작업 디렉터리를 루트 디렉터리로 설정하는 자동 폴백이고, 둘째 방법은 초기 작업 디렉터리를 변경할 수 있는 `pig.initial.fs.name` 속성입니다.

Amazon EMR 버전 5.0.0부터는 이 두 가지 방법을 사용할 수 없으며 사용자는 HDFS에 홈 디렉터리가 있어야 합니다. hadoop 사용자에게는 시작 시 홈 디렉터리가 할당되기 때문에 적용되지 않습니다. `command-runner.jar`를 사용하여 다른 사용자가 명시적으로 지정되지 않는 한, 하둡 jar 단계를 사용하여 스크립트를 실행하면 기본적으로 하둡 사용자로 지정됩니다.

Amazon EMR 릴리스 4.9.6

4.9.6 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, PhoenixPig, Oozie-SandboxPresto-Sandbox, Spark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-4.9.6	emr-4.9.5	emr-4.9.4	emr-4.9.3
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

4.9.6 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

4.9.6 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.17.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-9	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-9	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-9	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-9	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-9	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스

구성 요소	버전	설명
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.

구성 요소	버전	설명
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.9.6 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-4.9.6 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.

분류	설명
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.

분류	설명
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.9.5

4.9.5 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.9.5	emr-4.9.4	emr-4.9.3	emr-4.9.2
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

4.9.5 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.5에 대한 정보가 포함됩니다. 변경 사항은 4.9.4에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- HBase
  - 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

4.9.5 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.17.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-9	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-9	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-9	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	1.0.0-amzn-9	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-9	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.9.5 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.9.5 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.

분류	설명
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.9.4

4.9.4 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, PhoenixPig, Oozie-SandboxPresto-Sandbox, Spark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.9.4	emr-4.9.3	emr-4.9.2	emr-4.9.1
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

4.9.4 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.4에 대한 정보가 포함됩니다. 변경 사항은 4.9.3에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

4.9.4 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.17.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-9	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-9	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-9	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	1.0.0-amzn-9	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-9	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.9.4 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.9.4 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.

분류	설명
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.9.3

4.9.3 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Oozie-SandboxPhoenix, Presto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.9.3	emr-4.9.2	emr-4.9.1	emr-4.8.5
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

4.9.3 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.3에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.9.2 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

4.9.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.17.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-9	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-9	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-9	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-9	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-9	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.9.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.9.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.9.2

4.9.2 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Oozie-SandboxPhoenix, Presto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1

	emr-4.9.2	emr-4.9.1	emr-4.8.5	emr-4.8.4
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

4.9.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.2에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.9.1 릴리스에 관련됩니다.

릴리스 날짜: 2017년 7월 13일

이번 릴리스에서 일반적인 변경, 버그 수정 및 기능 개선이 이루어졌습니다.

4.9.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.3.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리

구성 요소	버전	설명
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.17.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-9	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	1.0.0-amzn-9	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog에 REST 인터페이스 를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-9	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-9	하둡 작업에 대한 SQL용 메타 데이터가 저장되는 시맨틱 리 포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-9	Hive 쿼리를 웹 요청으로 수락 하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션 을 사용하여 데이터를 분석하 는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하 는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세 스를 비롯하여 JDBC 액세스를 제공하는 경량 서버

구성 요소	버전	설명
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.9.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](https://docs.aws.amazon.com/emr/latest/ReleaseGuide/emr-configure-apps.html) 단원을 참조하십시오.

emr-4.9.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.

분류	설명
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.

분류	설명
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.9.1

4.9.1 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.157.1
Spark	1.6.3	1.6.3	1.6.3	1.6.3
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1

	emr-4.9.1	emr-4.8.5	emr-4.8.4	emr-4.8.3
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.9

4.9.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.4 릴리스에 관련됩니다.

릴리스 날짜: 2017년 4월 10일

이전 릴리스에서 해결된 알려진 문제

- HIVE-9976 및 HIVE-10106 백포트
- YARN에서 다수의 노드(2,000개 이상) 및 컨테이너(5,000개 이상)가 메모리 부족 오류를 야기하던 문제(예: `"Exception in thread main java.lang.OutOfMemoryError"`)가 수정되었습니다.

변경 사항 및 기능 향상

- Amazon EMR 릴리스는 이제 Amazon Linux 2017.03을 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/ 단원을 참조하십시오.
- Amazon EMR 기본 Linux 이미지에서 Python 2.6을 제거했습니다. 필요하다면 Python 2.6을 수동으로 설치할 수도 있습니다.

4.9.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서

시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.3.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.15.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.7.3-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-9	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-9	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-9	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-9	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-9	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-9	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스

구성 요소	버전	설명
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.9.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.9.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.

분류	설명
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.

분류	설명
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.8.5

4.8.5 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.3
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.157.1	0.152.3
Spark	1.6.3	1.6.3	1.6.3	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.8.5	emr-4.8.4	emr-4.8.3	emr-4.8.2
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.9	3.4.8

4.8.5 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

4.8.5 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.14.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-8	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-8	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-8	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-8	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-8	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-8	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트

구성 요소	버전	설명
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.25+	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.8.5 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.8.5 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.

분류	설명
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.8.4

4.8.4 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.3	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.157.1	0.152.3	0.151

	emr-4.8.4	emr-4.8.3	emr-4.8.2	emr-4.8.1
Spark	1.6.3	1.6.3	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.9	3.4.8	3.4.8

4.8.4 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.8.4에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.3 릴리스에 관련됩니다.

릴리스 날짜: 2017년 2월 7일

이번 릴리스에서 일반적인 변경, 버그 수정 및 기능 개선이 이루어졌습니다.

4.8.4 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이

터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.14.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-8	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-8	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-8	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-8	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-8	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-8	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.54+	MySQL 데이터베이스 서버.

구성 요소	버전	설명
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.25+	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.8.4 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.8.4 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.

분류	설명
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.

분류	설명
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.8.3

4.8.3 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.2
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.3	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-4.8.3	emr-4.8.2	emr-4.8.1	emr-4.8.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.157.1	0.152.3	0.151	0.151
Spark	1.6.3	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.4
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.9	3.4.8	3.4.8	3.4.8

4.8.3 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.8.3에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.2 릴리스에 관련됩니다.

릴리스 날짜: 2016년 12월 29일

업그레이드

- Presto 0.157.1로 업그레이드되었습니다. 자세한 내용은 Presto 설명서의 Presto 릴리스 정보를 참조하십시오.
- Spark 1.6.3으로 업그레이드되었습니다. 자세한 내용은 Apache Spark 설명서의 Spark 릴리스 정보를 참조하십시오.
- ZooKeeper 3.4.9로 업그레이드되었습니다. 자세한 내용은 Apache ZooKeeper 설명서의 ZooKeeper 릴리스 정보를 참조하십시오.

변경 사항 및 기능 향상

- Amazon EMR 버전 4.8.3 이상(5.0.0, 5.0.3 및 5.2.0 제외)에서 Amazon EC2 m4.16xlarge 인스턴스 유형에 대한 지원을 추가했습니다.
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.

이전 릴리스에서 해결된 알려진 문제

- 대규모 클러스터에서 동일한 파일의 복제 및 삭제 간 경합 때문에 ReplicationMonitor 스레드가 장시간 정체될 수 있는 하둡 문제를 수정했습니다.
- 작업 상태가 성공적으로 업데이트되지 않은 경우 ControlledJob#toString이 null 포인터 예외(NPE)와 함께 실패하는 문제를 수정했습니다.

4.8.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.2.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.13.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.3-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리

구성 요소	버전	설명
hadoop-hdfs-namenode	2.7.3-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-1	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스

구성 요소	버전	설명
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-8	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-8	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-8	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-8	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-8	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-8	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리

구성 요소	버전	설명
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.157.1	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.157.1	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.3	Spark 명령줄 클라이언트.
spark-history-server	1.6.3	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.3	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.3	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북

구성 요소	버전	설명
zookeeper-server	3.4.9	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.9	ZooKeeper 명령줄 클라이언트.

4.8.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.8.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.

분류	설명
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.

분류	설명
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.

분류	설명
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.8.2

4.8.2 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.2	1.2.1

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.3	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.152.3	0.151	0.151	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6

	emr-4.8.2	emr-4.8.1	emr-4.8.0	emr-4.7.4
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.4	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.6.1	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

4.8.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.8.2에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 10월 24일

업그레이드

- Hadoop 2.7.3로 업그레이드
- Presto 0.152.3로 업그레이드했습니다. 여기에는 Presto 웹 인터페이스에 대한 지원이 포함됩니다. 포트 8889를 사용하여 Presto 코디네이터의 Presto 웹 인터페이스에 액세스할 수 있습니다. Presto 웹 인터페이스에 대한 자세한 내용은 Presto 설명서의 웹 인터페이스를 참조하십시오.
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.

4.8.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을

위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	4.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.10.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.7.3-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.3-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.3-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.3-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.3-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.3-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.3-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.3-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.3-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.3-amzn-0	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-7	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-7	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-7	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-7	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-7	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-7	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.52	MySQL 데이터베이스 서버.

구성 요소	버전	설명
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.152.3	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.152.3	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.2	Spark 명령줄 클라이언트.
spark-history-server	1.6.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.8.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.8.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.

분류	설명
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.

분류	설명
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.

분류	설명
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.8.1

4.8.1 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.2	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0

	emr-4.8.1	emr-4.8.0	emr-4.7.4	emr-4.7.3
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.151	0.148	0.148
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.4	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.6.1	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

4.8.1 릴리스 정보

4.8.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.9.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.7.2-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스

구성 요소	버전	설명
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-7	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-7	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-7	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-7	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-7	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-7	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.51	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트

구성 요소	버전	설명
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.151	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.151	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.2	Spark 명령줄 클라이언트.
spark-history-server	1.6.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.

구성 요소	버전	설명
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.8.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.8.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.

분류	설명
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.

분류	설명
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.8.0

4.8.0 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.2	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.2
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.151	0.148	0.148	0.148

	emr-4.8.0	emr-4.7.4	emr-4.7.3	emr-4.7.2
Spark	1.6.2	1.6.2	1.6.2	1.6.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.4	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.6.1	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

4.8.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.8.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.7.2 릴리스에 관련됩니다.

릴리스 날짜: 2016년 9월 7일

업그레이드

- HBase 1.2.2로 업그레이드
- Presto-Sandbox 0.151으로 업그레이드
- Tez 0.8.4로 업그레이드
- Zeppelin-Sandbox 0.6.1으로 업그레이드

변경 사항 및 기능 향상

- YARN에서 인스턴스가 종료되었기 때문에 더 이상 존재하지 않는 컨테이너를 ApplicationMaster가 정리하려고 시도하는 문제를 수정했습니다.
- Oozie 예제에서 Hive2 작업에 대한 hive-server2 URL을 수정했습니다.
- 추가 Presto 카탈로그에 대한 지원을 추가했습니다.
- 백포트된 패치: HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- 암호화 옵션을 더 쉽게 생성하고 적용할 수 있도록 보안 구성에 대한 지원을 추가했습니다. 자세한 내용은 데이터 암호화를 참조하십시오.

4.8.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.9.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-4	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-4	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-4	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-4	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-4	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-4	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-4	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.2-amzn-4	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-4	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-4	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.2	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.2	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.2	HBase 명령줄 클라이언트
hbase-rest-server	1.2.2	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.2	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-7	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-7	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-7	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-7	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	1.0.0-amzn-7	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-7	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.51	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.151	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.151	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.2	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	1.6.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.4	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.6.1	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.8.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.8.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hiveserver2-site	Hive Server2 hiveserver2-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.

분류	설명
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-blackhole	Presto blackhole.properties 파일의 값을 변경합니다.

분류	설명
presto-connector-cassandra	Presto cassandra.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
presto-connector-jmx	Presto jmx.properties 파일의 값을 변경합니다.
presto-connector-kafka	Presto kafka.properties 파일의 값을 변경합니다.
presto-connector-localfile	Presto localfile.properties 파일의 값을 변경합니다.
presto-connector-mongodb	Presto mongodb.properties 파일의 값을 변경합니다.
presto-connector-mysql	Presto mysql.properties 파일의 값을 변경합니다.
presto-connector-postgresql	Presto postgresql.properties 파일의 값을 변경합니다.
presto-connector-raptor	Presto raptor.properties 파일의 값을 변경합니다.
presto-connector-redis	Presto redis.properties 파일의 값을 변경합니다.
presto-connector-tpch	Presto tpch.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.7.4

4.7.4 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.7.4	emr-4.7.3	emr-4.7.2	emr-4.7.1
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.2	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.148	0.147
Spark	1.6.2	1.6.2	1.6.2	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

4.7.4 릴리스 정보

Amazon S3에 대한 요청에 AWS 서명 버전 4 인증을 추가하는 패치 릴리스입니다. 모든 애플리케이션 및 구성 요소는 이전 Amazon EMR 릴리스와 동일합니다.

> **Important**
>
> 이 릴리스 버전에서 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 자세한 내용은 새 소식을 참조하십시오.

4.7.4 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.8.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.2-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.1	HBase 명령줄 클라이언트
hbase-rest-server	1.2.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-6	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-6	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-6	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-6	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	1.0.0-amzn-6	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-6	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.148	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.148	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.2	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	1.6.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.3	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.7.4 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-4.7.4 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.7.3

4.7.3 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Oozie-SandboxPhoenix, PigPresto-Sandbox, Spark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.75
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.1
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.2	0.12.0	0.12.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	4.7.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.148	0.147	0.147
Spark	1.6.2	1.6.2	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	0.8.3
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6

	emr-4.7.3	emr-4.7.2	emr-4.7.1	emr-4.7.0
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

4.7.3 릴리스 정보

4.7.3 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.8.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.2-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.1	HBase 명령줄 클라이언트
hbase-rest-server	1.2.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-6	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-6	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-6	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-6	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	1.0.0-amzn-6	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-6	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.148	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.148	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.2	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	1.6.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.2	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.3	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.7.3 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.7.3 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.

분류	설명
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.

분류	설명
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.

분류	설명
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.7.2

4.7.2 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
AWS Java용 SDK	1.10.75	1.10.75	1.10.75	1.10.27
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.1	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.2	0.12.0	0.12.0	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	4.7.0	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.148	0.147	0.147	0.143
Spark	1.6.2	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	0.8.3	-
Trino (PrestoSQ L)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6

	emr-4.7.2	emr-4.7.1	emr-4.7.0	emr-4.6.1
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

4.7.2 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.7.2에 대한 정보가 포함됩니다.

릴리스 날짜: 2016년 7월 15일

Features

- Mahout 0.12.2로 업그레이드
- Presto 0.148으로 업그레이드
- Spark 1.6.2로 업그레이드
- 이제 URI를 파라미터로 사용하여 EMRFS과 함께 사용할 AWSCredentialsProvider를 생성할 수 있습니다. 자세한 내용은 EMRFS에 대한 AWSCredentialsProvider 생성을 참조하십시오.
- 이제 EMRFS에서 사용자는 `emrfs-site.xml`의 `fs.s3.consistent.dynamodb.endpoint` 속성을 사용하여 일관된 보기 메타데이터에 대한 사용자 지정 DynamoDB 엔드포인트를 구성할 수 있습니다.
- `/usr/bin`이라고 하는 `spark-example`에 스크립트가 추가되면서 `/usr/lib/spark/spark/bin/run-example`을 래핑하여 직접 예제를 실행할 수 있습니다. 예를 들어, Spark 배포와 함께 제공되는 SparkPi 예제를 실행하려면 명령줄에서 또는 `spark-example SparkPi 100`을 API에서 단계로 사용하여 `command-runner.jar`를 실행할 수 있습니다.

이전 릴리스에서 해결된 알려진 문제

- Oozie 와 함께 Spark도 설치된 경우 `spark-assembly.jar`이 올바른 위치에 있지 않아서 Spark 애플리케이션이 Oozie와 함께 시작되지 못하는 문제를 수정했습니다.
- YARN 컨테이너에서 Spark Log4j 기반 로깅 문제를 수정했습니다.

4.7.2 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.2.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.1.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.8.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-3	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-3	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-3	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-3	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-3	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-3	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-3	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-3	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-3	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-3	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스

구성 요소	버전	설명
hbase-hmaster	1.2.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.1	HBase 명령줄 클라이언트
hbase-rest-server	1.2.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-6	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-6	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-6	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-6	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-6	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-6	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스

구성 요소	버전	설명
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.2	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.148	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.148	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.2	Spark 명령줄 클라이언트.
spark-history-server	1.6.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.2	YARN용 인 메모리 실행 엔진.

구성 요소	버전	설명
spark-yarn-slave	1.6.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.3	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.23	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.7.2 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.7.2 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.

분류	설명
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hadoop-ssl-server	hadoop ssl 서버 구성을 변경합니다.
hadoop-ssl-client	hadoop ssl 클라이언트 구성을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.

분류	설명
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.

분류	설명
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.

분류	설명
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.7.1

4.7.1 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Phoenix, Oozie-SandboxPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
AWS Java용 SDK	1.10.75	1.10.75	1.10.27	1.10.27
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.1	1.2.0	1.2.0
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.7.1	emr-4.7.0	emr-4.6.1	emr-4.6.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.12.0	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	4.7.0	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.147	0.143	0.143
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	0.8.3	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	3.4.8

4.7.1 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.7.1에 대한 정보가 포함됩니다.

릴리스 날짜: 2016년 6월 10일

이전 릴리스에서 해결된 알려진 문제

- VPC에서 프라이빗 서브넷을 사용하여 시작된 클러스터의 스타트업 시간을 연장하는 문제를 수정했습니다. 이 버그는 Amazon EMR 4.7.0 릴리스를 사용하여 시작된 클러스터에만 영향을 미칩니다.
- Amazon EMR 4.7.0 릴리스를 사용하여 시작된 클러스터의 경우 Amazon EMR에서 파일 목록을 부적절하게 처리하는 문제를 수정했습니다.

4.7.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.

구성 요소	버전	설명
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분산 복사 애플리케이션.
emrfs	2.7.1	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버

구성 요소	버전	설명
hadoop-mapred	2.7.2-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.1	HBase 명령줄 클라이언트
hbase-rest-server	1.2.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-5	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-5	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	1.0.0-amzn-5	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-5	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-5	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-5	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.147	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.

구성 요소	버전	설명
presto-worker	0.147	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.1	Spark 명령줄 클라이언트.
spark-history-server	1.6.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.3	tez YARN 애플리케이션 및 라이브러리.
webserver	2.4.18	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.7.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.7.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.

분류	설명
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.

분류	설명
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.

분류	설명
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.7.0

4.7.0 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, Hive, , Hue, Mahout, Oozie-SandboxPhoenix, Presto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Tez Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
AWS Java용 SDK	1.10.75	1.10.27	1.10.27	1.10.27
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.1	1.2.0	1.2.0	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.2
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.12.0	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	4.7.0	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.147	0.143	0.143	0.140
Spark	1.6.1	1.6.1	1.6.1	1.6.1
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	0.8.3	-	-	-
Trino (PrestoSQL)	-	-	-	-

	emr-4.7.0	emr-4.6.1	emr-4.6.0	emr-4.5.0
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	3.4.8	-

4.7.0 릴리스 정보

> **Important**
>
> Amazon EMR 4.7.0은 더 이상 사용되지 않습니다. 대신 Amazon EMR 4.7.1 이상을 사용합니다.

릴리스 날짜: 2016년 6월 2일

Features

- Apache Phoenix 4.7.0 추가
- Apache Tez 0.8.3 추가
- HBase 1.2.1으로 업그레이드
- Mahout 0.12.0로 업그레이드
- Presto 0.147으로 업그레이드
- 를 1.10.75 AWS SDK for Java 로 업그레이드
- 사용자가 로컬 모드에서 Pig를 실행할 수 있도록 `mapreduce.cluster.local.dir`의 `mapred-site.xml` 속성에서 최종 플래그가 제거되었습니다.
- 클러스터에서 사용할 수 있는 Amazon Redshift JDBC 드라이버

  이제 Amazon Redshift JDBC 드라이버는 `/usr/share/aws/redshift/jdbc`에 포함됩니다. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar`은 JDBC 4.1 호환 Amazon Redshift 드라이버이고 `/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar`은 JDBC 4.0 호환

Amazon Redshift 드라이버입니다. 자세한 내용은 Amazon Redshift 관리 안내서에서 JDBC 연결 구성을 참조하세요.

- Java 8

  Presto를 제외하고 모든 애플리케이션에 사용되는 기본 JDK는 OpenJDK 1.7입니다. 하지만 OpenJDK 1.7 및 1.8이 모두 설치됩니다. 애플리케이션에 대해 `JAVA_HOME`을 설정하는 방법에 대한 자세한 내용은 Java 8을 사용하도록 애플리케이션 구성을 참조하십시오.

이전 릴리스에서 해결된 알려진 문제

- emr-4.6.0에서 Amazon EMR에 사용되는 처리량에 최적화된 HDD(st1) EBS 볼륨의 성능에 상당한 영향을 미치는 커널 문제를 수정했습니다.
- 하둡을 애플리케이션으로 선택하지 않고 HDFS 암호화 영역을 지정할 경우 클러스터가 실패하는 문제를 수정했습니다.
- 기본 HDFS 쓰기 정책을 `RoundRobin`에서 `AvailableSpaceVolumeChoosingPolicy`로 변경했습니다. RoundRobin 구성에서 일부 볼륨이 제대로 이용되지 않았으며 결과적으로 코어 노드가 실패하고 HDFS가 불안정한 상태가 되었습니다.
- 일관된 보기에 대한 기본 DynamoDB 메타데이터 테이블을 생성할 때 예외가 발생하는 EMRFS CLI 문제를 수정했습니다.
- 멀티파트 이름 바꾸기 및 복사 작업 중에 발생할 수 있는 EMRFS의 교착 문제를 수정했습니다.
- CopyPart 크기가 기본적으로 5MB로 설정되는 EMRFS 문제를 수정했습니다. 기본값은 이제 128MB로 올바르게 설정됩니다.
- 서비스를 중단할 수 없는 Zeppelin upstart 구성 문제를 수정했습니다.
- `s3a://`이 해당 클래스 경로에 올바르게 로드되지 않기 때문에 `/usr/lib/hadoop/hadoop-aws.jar` URI 스키마를 사용할 수 없는 Spark 및 Zeppelin 문제를 수정했습니다.
- HUE-2484를 백포트했습니다.
- HBase 브라우저 샘플 문제를 수정하기 위해 Hue 3.9.0(JIRA가 존재하지 않음)의 커밋을 백포트했습니다.
- HIVE-9073를 백포트했습니다.

4.7.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을

위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `CommunityVersion-amzn-EmrVersion` 양식의 버전 레이블이 있습니다. `EmrVersion`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.1.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.2.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.4.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.7.1	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.

구성 요소	버전	설명
hadoop-client	2.7.2-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hadoop-yarn-timeline-server	2.7.2-amzn-2	YARN 애플리케이션의 현재 및 기록 정보를 가져오는 서비스
hbase-hmaster	1.2.1	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스

구성 요소	버전	설명
hbase-region-server	1.2.1	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.1	HBase 명령줄 클라이언트
hbase-rest-server	1.2.1	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.1	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-5	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-5	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-5	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-5	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-5	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-5	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-7	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.12.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5.46	MySQL 데이터베이스 서버.

구성 요소	버전	설명
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
phoenix-library	4.7.0-HBase-1.2	서버 및 클라이언트의 phoenix 라이브러리
phoenix-query-server	4.7.0-HBase-1.2	Avatica API에 대한 JSON 형식 액세스 및 프로토콜 버퍼 액세스를 비롯하여 JDBC 액세스를 제공하는 경량 서버
presto-coordinator	0.147	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.147	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.1	Spark 명령줄 클라이언트.
spark-history-server	1.6.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
tez-on-yarn	0.8.3	tez YARN 애플리케이션 및 라이브러리.

구성 요소	버전	설명
webserver	2.4.18	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.7.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.7.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.

분류	설명
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.

분류	설명
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
phoenix-hbase-metrics	Phoenix hadoop-metrics2-hbase.properties 파일의 값을 변경합니다.
phoenix-hbase-site	Phoenix hbase-site.xml 파일의 값을 변경합니다.
phoenix-log4j	Phoenix log4j.properties 파일의 값을 변경합니다.

분류	설명
phoenix-metrics	Phoenix hadoop-metrics2-phoenix.properties 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
tez-site	Tez tez-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.

분류	설명
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.6.1

4.6.1 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Oozie-Sandbox, MahoutPresto-Sandbox, PigSpark, 및 애플리케이션이 포함되어 Sqoop-Sandbox Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
AWS Java용 SDK	1.10.27	1.10.27	1.10.27	1.10.27
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	1.2.0	-	-
HCatalog	1.0.0	1.0.0	1.0.0	1.0.0
Hadoop	2.7.2	2.7.2	2.7.2	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.1	0.11.1
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-

	emr-4.6.1	emr-4.6.0	emr-4.5.0	emr-4.4.0
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.143	0.140	0.136
Spark	1.6.1	1.6.1	1.6.1	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	1.4.6
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.6
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	3.4.8	-	-

4.6.1 릴리스 정보

4.6.1 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.1.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.3.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.6.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트

구성 요소	버전	설명
hadoop-hdfs-datanode	2.7.2-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hbase-hmaster	1.2.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.0	HBase 명령줄 클라이언트

구성 요소	버전	설명
hbase-rest-server	1.2.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-4	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-6	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.11.1	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스

구성 요소	버전	설명
presto-coordinator	0.143	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.143	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.1	Spark 명령줄 클라이언트.
spark-history-server	1.6.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
webserver	2.4	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.6.1 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.6.1 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.

분류	설명
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.

분류	설명
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.

분류	설명
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.6.0

4.6.0 애플리케이션 버전

이 릴리스에는 Ganglia, , HBase, HCatalog, Hadoop, , Hive, Hue, Oozie-Sandbox, MahoutPresto-Sandbox, PigSpark, Sqoop-Sandbox, 및 애플리케이션이 포함되어 Zeppelin-Sandbox있습니다ZooKeeper-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
AWS Java용 SDK	1.10.27	1.10.27	1.10.27	1.10.27
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.7.2
HBase	1.2.0	-	-	-
HCatalog	1.0.0	1.0.0	1.0.0	-
Hadoop	2.7.2	2.7.2	2.7.1	2.7.1
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.6.0	emr-4.5.0	emr-4.4.0	emr-4.3.0
Mahout	0.11.1	0.11.1	0.11.1	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.143	0.140	0.136	0.130
Spark	1.6.1	1.6.1	1.6.0	1.6.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	1.4.6	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.6	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	3.4.8	-	-	-

4.6.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.6.0에 대한 정보가 포함됩니다.

- HBase 1.2.0 추가
- Zookeeper-Sandbox 3.4.8 추가
- Presto-Sandbox 0.143로 업그레이드
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.03.0을 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/ 단원을 참조하십시오.
- 처리량에 최적화된 HDD(st1) EBS 볼륨 유형에 영향을 미치는 문제

  Linux 커널 버전 4.2 이상의 문제는 EMR에 사용되는 처리량에 최적화된 HDD(st1) EBS 볼륨의 성능에 상당한 영향을 미칩니다. 이 릴리스(emr-4.6.0)는 커널 버전 4.4.5를 사용하므로 영향을 받습니다. 따라서 st1 EBS 볼륨을 사용하려는 경우 emr-4.6.0를 사용하지 않는 것이 좋습니다. emr-4.5.0 이전의 Amazon EMR 릴리스를 st1과 함께 사용하면 영향을 받지 않을 수 있습니다. 또한 향후 릴리스에서 수정을 제공합니다.
- Python 기본값

  이제 Python 3.4가 기본적으로 설치되지만, 시스템 기본값은 여전히 Python 2.7입니다. 부트스트랩 작업을 사용하여 Python 3.4를 시스템 기본값으로 구성할 수 있습니다. 구성 API를 사용하여 `/usr/bin/python3.4` 분류의 PYSPARK_PYTHON 내보내기를 `spark-env`로 설정하면 PySpark에서 사용되는 Python 버전에 영향을 미칠 수 있습니다.
- Java 8

  Presto를 제외하고 모든 애플리케이션에 사용되는 기본 JDK는 OpenJDK 1.7입니다. 하지만 OpenJDK 1.7 및 1.8이 모두 설치됩니다. 애플리케이션에 대해 `JAVA_HOME`을 설정하는 방법에 대한 자세한 내용은 Java 8을 사용하도록 애플리케이션 구성을 참조하십시오.

이전 릴리스에서 해결된 알려진 문제

- 생성된 암호로 인해 애플리케이션 프로비저닝이 가끔 임의로 실패하는 문제를 수정했습니다.
- 이전에는 `mysqld`가 모든 노드에 설치되었습니다. 이제 이 항목은 마스터 인스턴스에만 설치되며 선택한 애플리케이션에 `mysql-server`가 구성 요소로 포함되는 경우에만 설치됩니다. 현재 HCatalog, Hive, Hue, Presto-Sandbox, Sqoop-Sandbox 등의 애플리케이션에 `mysql-server` 구성 요소가 포함됩니다.

- `yarn.scheduler.maximum-allocation-vcores`를 기본값 32에서 80으로 변경했습니다. 이 변경에 따라 emr-4.4.0에 도입된 한 가지 문제가 수정됩니다. 이 문제는 코어 인스턴스 유형이 32 이상의 YARN vcore 세트를 갖춘 몇 가지 대규모 인스턴스 유형 중 하나인 클러스터에서 `maximizeResourceAllocation` 옵션을 사용하는 동안 주로 Spark에 발생합니다. 다시 말해서 c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge 또는 m4.10xlarge가 이 문제의 영향을 받습니다.
- s3-dist-cp는 이제 모든 Amazon S3 추천에 EMRFS를 사용하며 더 이상 임시 HDFS 디렉터리로 준비하지 않습니다.
- 클라이언트 측 암호화 멀티파트 업로드에 대한 예외 처리 문제를 수정했습니다.
- 사용자가 Amazon S3 스토리지 클래스를 변경할 수 있는 옵션을 추가했습니다. 기본적으로 이 설정은 `STANDARD`입니다. `emrfs-site` 구성 분류 설정은 `fs.s3.storageClass`이며 가능한 값은 `STANDARD`, `STANDARD_IA` 및 `REDUCED_REDUNDANCY`입니다. 스토리지 클래스에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 [스토리지 클래스]를 참조하세요.

4.6.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리

구성 요소	버전	설명
emr-kinesis	3.1.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.3.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.6.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트

구성 요소	버전	설명
hadoop-kms-server	2.7.2-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.2-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hbase-hmaster	1.2.0	리전 조정 및 관리 명령의 실행을 담당하는 HBase 클러스터용 서비스
hbase-region-server	1.2.0	한 개 이상의 HBase 리전에 서비스를 제공하는 서비스
hbase-client	1.2.0	HBase 명령줄 클라이언트
hbase-rest-server	1.2.0	HBase용 RESTful HTTP 엔드포인트를 제공하는 서비스
hbase-thrift-server	1.2.0	HBase에 Thrift 엔드포인트를 제공하는 서비스
hcatalog-client	1.0.0-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스

구성 요소	버전	설명
hcatalog-webhcat-server	1.0.0-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-4	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-6	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.11.1	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
presto-coordinator	0.143	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.143	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.1	Spark 명령줄 클라이언트.

구성 요소	버전	설명
spark-history-server	1.6.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
webserver	2.4	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북
zookeeper-server	3.4.8	구성 정보 유지 관리, 이름 지정, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙 집중식 서비스.
zookeeper-client	3.4.8	ZooKeeper 명령줄 클라이언트.

4.6.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.6.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.

분류	설명
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hbase-env	HBase 환경의 값을 변경합니다.
hbase-log4j	HBase hbase-log4j.properties 파일의 값을 변경합니다.
hbase-metrics	HBase hadoop-metrics2-hbaase.properties 파일의 값을 변경합니다.
hbase-policy	HBase hbase-policy.xml 파일의 값을 변경합니다.
hbase-site	HBase hbase-site.xml 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.

분류	설명
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

분류	설명
zookeeper-config	ZooKeeper zoo.cfg 파일의 값을 변경합니다.
zookeeper-log4j	ZooKeeper log4j.properties 파일의 값을 변경합니다.

Amazon EMR 릴리스 4.5.0

4.5.0 애플리케이션 버전

이 릴리스에는 Ganglia, , HCatalog, Hadoop, Hive, , Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark및 애플리케이션이 포함되어 Sqoop-Sandbox있습니다Zeppelin-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
AWS Java용 SDK	1.10.27	1.10.27	1.10.27	1.10.27
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.7.2	3.6.0
HBase	-	-	-	-
HCatalog	1.0.0	1.0.0	-	-
Hadoop	2.7.2	2.7.1	2.7.1	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.1	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.2.0
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-

	emr-4.5.0	emr-4.4.0	emr-4.3.0	emr-4.2.0
Presto-Sandbox	0.140	0.136	0.130	0.125
Spark	1.6.1	1.6.0	1.6.0	1.5.2
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	1.4.6	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.6	0.5.5	0.5.5
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

4.5.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.5.0에 대한 정보가 포함됩니다.

릴리스 날짜: 2016년 4월 4일

Features

- Spark 1.6.1으로 업그레이드
- Hadoop 2.7.2로 업그레이드
- Presto 0.140로 업그레이드
- Amazon S3 서버 측 암호화에 대한 AWS KMS 지원이 추가되었습니다.

이전 릴리스에서 해결된 알려진 문제

- 노드를 재부팅한 후 MySQL 및 Apache 서버가 시작되지 않는 문제를 수정했습니다.
- Amazon S3에 파티션되지 않은 테이블이 저장되면 IMPORT가 올바르게 작동하지 않는 문제를 수정했습니다.
- Hive 테이블에 쓸 때 준비 디렉터리를 `/mnt/tmp` 대신 `/tmp`로 지정해야 하는 Presto 문제를 수정했습니다.

4.5.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 `*CommunityVersion*-amzn-*EmrVersion*` 양식의 버전 레이블이 있습니다. `*EmrVersion*`은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.1.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.2.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.5.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.

구성 요소	버전	설명
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.2-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.2-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.2-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.2-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.2-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.2-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.2-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.2-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스

구성 요소	버전	설명
hadoop-yarn-resourcemanager	2.7.2-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hcatalog-client	1.0.0-amzn-4	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트
hcatalog-server	1.0.0-amzn-4	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-4	HCatalog에 REST 인터페이스를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-4	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-4	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-4	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-5	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.11.1	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스

구성 요소	버전	설명
presto-coordinator	0.140	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.140	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.6.1	Spark 명령줄 클라이언트.
spark-history-server	1.6.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
webserver	2.4	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북

4.5.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

emr-4.5.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

Amazon EMR 릴리스 4.4.0

4.4.0 애플리케이션 버전

이 릴리스에는 Ganglia, , HCatalog, Hadoop, Hive, , Hue, Mahout, Oozie-SandboxPig, Presto-Sandbox, Spark및 애플리케이션이 포함되어 Sqoop-Sandbox있습니다Zeppelin-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
AWS Java용 SDK	1.10.27	1.10.27	1.10.27	추적되지 않음
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.7.2	3.6.0	-
HBase	-	-	-	-
HCatalog	1.0.0	-	-	-
Hadoop	2.7.1	2.7.1	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
Hue	3.7.1	3.7.1	3.7.1	3.7.1
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.1	0.11.0	0.11.0	0.11.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.2.0	4.0.1
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.136	0.130	0.125	0.119
Spark	1.6.0	1.6.0	1.5.2	1.5.0
Sqoop	-	-	-	-
Sqoop-Sandbox	1.4.6	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-

	emr-4.4.0	emr-4.3.0	emr-4.2.0	emr-4.1.0
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.6	0.5.5	0.5.5	0.6.0-SNAPSHOT
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

4.4.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.4.0에 대한 정보가 포함됩니다.

릴리스 날짜: 2016년 3월 14일

Features

- HCatalog 1.0.0 추가
- Sqoop-Sandbox 1.4.6 추가
- Presto 0.136로 업그레이드
- Zeppelin 0.5.6로 업그레이드
- Mahout 0.11.1으로 업그레이드
- `dynamicResourceAllocation`을 기본적으로 활성화했습니다.
- 릴리스에 대한 모든 구성 분류의 표를 추가했습니다. 자세한 내용은 애플리케이션 구성의 구성 분류 표를 참조하십시오.

이전 릴리스에서 해결된 알려진 문제

- `maximizeResourceAllocation` 설정에서 YARN ApplicationMaster 데몬에 충분한 메모리가 예약되지 않는 문제를 수정했습니다.
- 사용자 지정 DNS에 발생하는 문제를 수정했습니다. 제공된 사용자 지정 항목 앞에 `resolve.conf`의 항목이 있으면 사용자 지정 항목을 확인할 수 없습니다. 이 동작은 기본 VPC 이름 서버가 `resolve.conf`에 최상위 항목으로 삽입되는 VPC의 클러스터에서 영향을 받은 것입니다.

- 기본 Python이 버전 2.7으로 이동하고 해당 버전에 대해 boto가 설치되지 않는 문제를 수정했습니다.
- YARN 컨테이너와 Spark 애플리케이션이 고유의 Ganglia rrd(round robin database) 파일을 생성하여 인스턴스에 연결된 첫 번째 디스크가 가득 채워지는 문제를 수정했습니다. 이러한 수정으로 인해 YARN 컨테이너 수준 지표가 비활성화되었으며 Spark 애플리케이션 수준 지표가 비활성화되었습니다.
- 로그 푸셔에서 비어 있는 모든 로그 폴더가 삭제되는 문제를 수정했습니다. 이 문제의 영향으로 로그 푸셔가 `user`에서 비어 있는 `/var/log/hive` 폴더를 제거하기 때문에 Hive CLI가 로그를 작성할 수 없었습니다.
- Hive 가져오기에 영향을 미치는 문제를 수정했습니다. 이 문제는 파티셔닝에 영향을 미치며 결과적으로 가져오기 중 오류가 발생했습니다.
- EMRFS 및 s3-dist-cp에서 마침표가 포함된 버킷 이름이 올바르게 처리되지 않는 문제를 수정했습니다.
- 버전 관리를 사용하는 버킷에서 `_$folder$` 마커 파일이 연속적으로 생성되지 않도록 EMRFS의 동작을 변경했습니다. 이 변경은 버전 관리를 사용하는 버킷의 성능을 향상하는 데 도움이 될 수 있습니다.
- 클라이언트 측 암호화를 사용하는 경우를 제외하고 지침 파일을 사용하지 않도록 EMRFS의 동작을 변경했습니다. 클라이언트 측 암호화를 사용하는 동안 지침 파일을 삭제하려는 경우 emrfs-site.xml 속성인 `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled`를 true로 설정할 수 있습니다.
- 2일 동인 집계 대상에서 로그를 유지하도록 Changed YARN 로그 집계를 변경했습니다. 기본 대상은 클러스터의 HDFS 스토리지입니다. 이 기간을 변경하려는 경우 클러스터를 생성할 때 `yarn.log-aggregation.retain-seconds` 분류를 사용하여 `yarn-site`의 값을 변경합니다. 평소와 같이, 클러스터를 생성할 때 `log-uri` 파라미터를 사용하여 애플리케이션 로그를 Amazon S3에 저장할 수 있습니다.

적용된 패치

- HIVE-9655
- HIVE-9183
- HADOOP-12810

4.4.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.1.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.2.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.4.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.

구성 요소	버전	설명
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.1-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.1-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.1-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.1-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.1-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.1-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.1-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.7.1-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.1-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hcatalog-client	1.0.0-amzn-3	hcatalog-server를 조정하기 위한 'hcat' 명령줄 클라이언트

구성 요소	버전	설명
hcatalog-server	1.0.0-amzn-3	분산 애플리케이션용 스토리지 관리 계층, 테이블 및 HCatalog 를 제공하는 서비스
hcatalog-webhcat-server	1.0.0-amzn-3	HCatalog에 REST 인터페이스 를 제공하는 HTTP 엔드포인트
hive-client	1.0.0-amzn-3	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-3	하둡 작업에 대한 SQL용 메타 데이터가 저장되는 시맨틱 리 포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-3	Hive 쿼리를 웹 요청으로 수락 하기 위한 서비스
hue-server	3.7.1-amzn-5	하둡 에코시스템 애플리케이션 을 사용하여 데이터를 분석하 는 웹 애플리케이션.
mahout-client	0.11.1	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하 는 서비스
presto-coordinator	0.136	presto-worker의 쿼리를 수락하 고 쿼리 실행을 관리하는 서비 스.
presto-worker	0.136	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.

구성 요소	버전	설명
spark-client	1.6.0	Spark 명령줄 클라이언트.
spark-history-server	1.6.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
sqoop-client	1.4.6	Apache Sqoop 명령줄 클라이언트.
webserver	2.4	Apache HTTP 서버.
zeppelin-server	0.5.6-incubating	대화형 데이터 분석이 가능한 웹 기반 노트북

4.4.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.4.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.

분류	설명
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hcatalog-env	HCatalog 환경의 값을 변경합니다.
hcatalog-server-jndi	HCatalog jndi.properties 환경의 값을 변경합니다.
hcatalog-server-proto-hive-site	HCatalog proto-hive-site.xml의 값을 변경합니다.
hcatalog-webhcat-env	HCatalog WebHCat 환경의 값을 변경합니다.
hcatalog-webhcat-log4j	HCatalog WebHCat's log4j.properties의 값을 변경합니다.
hcatalog-webhcat-site	HCatalog WebHCat webhcat-site.xml 파일의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.

분류	설명
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.

분류	설명
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
sqoop-env	Sqoop 환경의 값을 변경합니다.
sqoop-oraoop-site	Sqoop OraOop oraoop-site.xml 파일의 값을 변경합니다.
sqoop-site	Sqoop sqoop-site.xml 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

Amazon EMR 릴리스 4.3.0

4.3.0 애플리케이션 버전

이 릴리스에는 Ganglia, , Hadoop, Hive, Hue, , Mahout, Oozie-SandboxPig, Presto-Sandbox및 애플리케이션이 포함되어 Spark있습니다Zeppelin-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS Java용 SDK	1.10.27	1.10.27	추적되지 않음	추적되지 않음
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

4.3.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.3.0에 대한 정보가 포함됩니다.

릴리스 날짜: 2016년 1월 19일

Features

- Hadoop 2.7.1으로 업그레이드
- Spark 1.6.0로 업그레이드
- Ganglia를 3.7.2로 업그레이드
- Presto를 0.130로 업그레이드
- Amazon EMR에서 `spark.dynamicAllocation.enabled`를 true로 설정한 경우 이 설정이 약간 변경되었습니다. 이 설정은 기본적으로 false입니다. 이 값을 true로 설정하면 이 설정은 `maximizeResourceAllocation` 설정을 통해 지정된 기본값에 영향을 미칩니다.
  - `spark.dynamicAllocation.enabled`가 true로 설정되면 `spark.executor.instances`에서 `maximizeResourceAllocation`가 설정되지 않습니다.
  - 이제 `spark.driver.memory` 설정은 `spark.executors.memory`가 설정되는 방식과 비슷한 방식으로 클러스터의 인스턴스 유형에 따라 구성됩니다. 하지만 Spark 드라이버 애플리케이션을 마스터 또는 코어 인스턴스 중 하나에서(예를 들면 각각 YARN 클라이언트 및 클러스터 모드에서) 실행할 수 있으므로 `spark.driver.memory` 설정이 이러한 두 인스턴스 그룹 중 더 작은 인스턴스 유형에 따라 설정됩니다.
  - 이제 `spark.default.parallelism` 설정은 YARN 컨테이너에 사용할 수 있는 CPU 코어 수의 두 배로 설정됩니다. 이전 릴리스에서 이 설정은 해당 값의 절반이었습니다.
  - Spark YARN 프로세스에서 예약된 메모리 오버헤드에 대한 계산이 더 정확하게 조정되어 Spark에 사용 가능한 총 메모리 양(즉, `spark.executor.memory`)이 약간 증가했습니다.

이전 릴리스에서 해결된 알려진 문제

- 이제 YARN 로그 집계가 기본적으로 활성화됩니다.
- YARN 로그 집계를 활성화하면 로그가 클러스터의 Amazon S3 로그 버킷에 푸시되지 않는 문제를 수정했습니다.
- 이제 YARN 컨테이너 크기의 새로운 최소값은 모든 노드 유형에서 32입니다.
- 대규모 클러스터의 프라이머리 노드에 과도한 디스크 I/O를 유발하는 Ganglia 문제를 수정했습니다.
- 클러스터를 종료할 때 애플리케이션 로그가 Amazon S3에 푸시되지 않는 문제를 수정했습니다.

- EMRFS CLI에서 특정 명령이 실패하는 문제를 수정했습니다.
- 종속성이 기본 SparkContext에 로드되지 않는 Zeppelin 문제를 수정했습니다.
- 크기 조정을 실행하여 인스턴스를 추가하려고 할 때 발생하는 문제를 수정했습니다.
- Hive에서 CREATE TABLE AS SELECT를 실행하면 Amazon S3에 대한 목록이 과도하게 직접 호출되는 문제를 수정했습니다.
- Hue, Oozie 및 Ganglia를 설치하면 대규모 클러스터가 제대로 프로비저닝되지 않는 문제를 수정했습니다.
- s3-dist-cp가 오류로 실패하더라도 0 종료 코드가 반환되는 문제를 수정했습니다.

4.3.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.1.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.1.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.

구성 요소	버전	설명
emrfs	2.3.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.7.2	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.
ganglia-metadata-collector	3.7.2	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.7.1	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.7.1-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.7.1-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.7.1-amzn-0	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.7.1-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.7.1-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.7.1-amzn-0	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.7.1-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리

구성 요소	버전	설명
hadoop-yarn-nodemanager	2.7.1-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.7.1-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hive-client	1.0.0-amzn-2	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-2	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-2	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-5	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.11.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
presto-coordinator	0.130	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.130	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.

구성 요소	버전	설명
spark-client	1.6.0	Spark 명령줄 클라이언트.
spark-history-server	1.6.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.6.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.6.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
webserver	2.4	Apache HTTP 서버.
zeppelin-server	0.5.5-incubating-amzn-1	대화형 데이터 분석이 가능한 웹 기반 노트북

4.3.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.3.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.

분류	설명
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.

분류	설명
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

Amazon EMR 릴리스 4.2.0

4.2.0 애플리케이션 버전

이 릴리스에는 Ganglia, , Hadoop, Hive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox및 애플리케이션이 포함되어 Spark있습니다Zeppelin-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS Java용 SDK	1.10.27	1.10.27	추적되지 않음	추적되지 않음
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQ L)	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

4.2.0 릴리스 정보

다음 릴리스 정보에는 Amazon EMR 릴리스 4.2.0에 대한 정보가 포함됩니다.

릴리스 날짜: 2015년 11월 18일

Features

- Ganglia 지원 추가
- Spark 1.5.2로 업그레이드
- Presto 0.125로 업그레이드
- Oozie를 4.2.0로 업그레이드
- Zeppelin을 0.5.5로 업그레이드
- 를 1.10.27 AWS SDK for Java 로 업그레이드했습니다.

이전 릴리스에서 해결된 알려진 문제

- 기본 메타데이터 테이블 이름이 사용되지 않는 EMRFS CLI 문제를 수정했습니다.
- Amazon S3에서 ORC 지원 테이블을 사용할 때 발생하는 문제를 수정했습니다.
- Spark 구성에서 Python 버전이 일치하지 않는 경우 발생하는 문제를 수정했습니다.
- VPC에서 클러스터의 DNS 문제 때문에 YARN 노드 상태가 보고되지 않는 문제를 수정했습니다.
- YARN에서 노드를 폐기하면 애플리케이션이 정지되거나 새 애플리케이션을 예약할 수 없게 되는 문제를 수정했습니다.
- 클러스터가 TIMED_OUT_STARTING 상태로 종료될 때 발생하는 문제를 수정했습니다.

- EMRFS Scala 종속성을 다른 빌드에 포함시킬 때 발생하는 문제를 수정했습니다. Scala 종속성이 제거되었습니다.

4.2.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.1.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.0.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.2.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
ganglia-monitor	3.6.0	하둡 에코시스템 애플리케이션용 내장형 Ganglia 에이전트와 Ganglia 모니터링 에이전트.

구성 요소	버전	설명
ganglia-metadata-collector	3.6.0	Ganglia 모니터링 에이전트의 측정치를 집계하기 위한 Ganglia 메타데이터 수집기.
ganglia-web	3.5.10	Ganglia 메타데이터 수집기에서 수집되는 측정치를 볼 수 있는 웹 애플리케이션.
hadoop-client	2.6.0-amzn-2	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.6.0-amzn-2	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-library	2.6.0-amzn-2	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.6.0-amzn-2	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.6.0-amzn-2	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.6.0-amzn-2	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.6.0-amzn-2	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.6.0-amzn-2	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.6.0-amzn-2	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스

구성 요소	버전	설명
hive-client	1.0.0-amzn-1	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-5	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.
mahout-client	0.11.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
oozie-client	4.2.0	Oozie 명령줄 클라이언트
oozie-server	4.2.0	Oozie 워크플로 요청을 수락하는 서비스
presto-coordinator	0.125	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.125	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.5.2	Spark 명령줄 클라이언트.
spark-history-server	1.5.2	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.5.2	YARN용 인 메모리 실행 엔진.

구성 요소	버전	설명
spark-yarn-slave	1.5.2	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
webserver	2.4	Apache HTTP 서버.
zeppelin-server	0.5.5-incubating-amzn-0	대화형 데이터 분석이 가능한 웹 기반 노트북

4.2.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.2.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.

분류	설명
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.

분류	설명
presto-config	Presto config.properties 파일의 값을 변경합니다.
presto-connector-hive	Presto hive.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
spark-metrics	Spark metrics.properties 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

Amazon EMR 릴리스 4.1.0

4.1.0 애플리케이션 버전

이 릴리스에는 Hadoop, , Hive, Hue, Mahout, Oozie-Sandbox, Pig, Presto-Sandbox및 애플리케이션이 포함되어 Spark있습니다Zeppelin-Sandbox.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전

- [Amazon EMR 5.x 릴리스의 애플리케이션 버전]
- [Amazon EMR 4.x 릴리스의 애플리케이션 버전]

애플리케이션 버전 정보

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS Java용 SDK	1.10.27	1.10.27	추적되지 않음	추적되지 않음
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEnterpriseGateway	-	-	-	-
JupyterHub	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

4.1.0 릴리스 정보

4.1.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *`CommunityVersion`*`-amzn-`*`EmrVersion`* 양식의 버전 레이블이 있습니다. *`EmrVersion`*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.1.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.0.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.1.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
hadoop-client	2.6.0-amzn-1	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.6.0-amzn-1	블록을 저장하는 HDFS 노드 수준 서비스

구성 요소	버전	설명
hadoop-hdfs-library	2.6.0-amzn-1	HDFS 명령줄 클라이언트 및 라이브러리
hadoop-hdfs-namenode	2.6.0-amzn-1	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.6.0-amzn-1	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-kms-server	2.6.0-amzn-1	하둡의 KeyProvider API를 기반으로 하는 암호화 키 관리 서버
hadoop-mapred	2.6.0-amzn-1	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.6.0-amzn-1	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.6.0-amzn-1	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hive-client	1.0.0-amzn-1	Hive 명령줄 클라이언트
hive-metastore-server	1.0.0-amzn-1	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-1	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
hue-server	3.7.1-amzn-4	하둡 에코시스템 애플리케이션을 사용하여 데이터를 분석하는 웹 애플리케이션.

구성 요소	버전	설명
mahout-client	0.11.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
oozie-client	4.0.1	Oozie 명령줄 클라이언트
oozie-server	4.0.1	Oozie 워크플로 요청을 수락하는 서비스
presto-coordinator	0.119	presto-worker의 쿼리를 수락하고 쿼리 실행을 관리하는 서비스.
presto-worker	0.119	여러 쿼리를 실행하는 서비스.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.5.0	Spark 명령줄 클라이언트.
spark-history-server	1.5.0	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.5.0	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.5.0	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.
zeppelin-server	0.6.0-incubating-SNAPSHOT	대화형 데이터 분석이 가능한 웹 기반 노트북

4.1.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.1.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hdfs-encryption-zones	HDFS 암호화 영역을 구성합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
hue-ini	Hue ini 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
hadoop-kms-acls	하둡 kms-acls.xml 파일의 값을 변경합니다.
hadoop-kms-env	하둡 KMS 환경의 값을 변경합니다.
hadoop-kms-log4j	하둡 kms-log4j.properties 파일의 값을 변경합니다.

분류	설명
hadoop-kms-site	하둡 kms-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
oozie-env	Oozie 환경의 값을 변경합니다.
oozie-log4j	Oozie oozie-log4j.properties 파일의 값을 변경합니다.
oozie-site	Oozie oozie-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
presto-log	Presto log.properties 파일의 값을 변경합니다.
presto-config	Presto config.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.
zeppelin-env	Zeppelin 환경의 값을 변경합니다.

Amazon EMR 릴리스 4.0.0

4.0.0 애플리케이션 버전

이 릴리스에는 Hadoop, , Hive, 및 애플리케이션이 포함되어 Mahout Pig있습니다Spark.

아래 테이블에는 이번 Amazon EMR 릴리스에서 사용할 수 있는 애플리케이션 버전과 이전 세 가지 Amazon EMR 릴리스(해당하는 경우)의 애플리케이션 버전이 나와 있습니다.

각 Amazon EMR 릴리스에서 애플리케이션 버전의 전체 기록은 다음 주제를 참조하세요.

- Amazon EMR 7.x 릴리스의 애플리케이션 버전
- Amazon EMR 6.x 릴리스의 애플리케이션 버전
- Amazon EMR 5.x 릴리스의 애플리케이션 버전
- Amazon EMR 4.x 릴리스의 애플리케이션 버전

애플리케이션 버전 정보

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
AWS Java용 SDK	1.10.27	1.10.27	추적되지 않음	추적되지 않음
Python	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
Scala	추적되지 않음	추적되지 않음	추적되지 않음	추적되지 않음
AmazonCloudWatchAgent	-	-	-	-
Delta	-	-	-	-
Flink	-	-	-	-
Ganglia	3.7.2	3.6.0	-	-
HBase	-	-	-	-
HCatalog	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Hadoop	2.7.1	2.6.0	2.6.0	2.6.0
Hive	1.0.0	1.0.0	1.0.0	1.0.0
Hudi	-	-	-	-
Hue	3.7.1	3.7.1	3.7.1	-
Iceberg	-	-	-	-
JupyterEn terpriseGateway	-	-	-	-
JupyterHub	-	-	-	-
Livy	-	-	-	-
MXNet	-	-	-	-
Mahout	0.11.0	0.11.0	0.11.0	0.10.0
Oozie	-	-	-	-
Oozie-Sandbox	4.2.0	4.2.0	4.0.1	-
Phoenix	-	-	-	-
Pig	0.14.0	0.14.0	0.14.0	0.14.0
Presto	-	-	-	-
Presto-Sandbox	0.130	0.125	0.119	-
Spark	1.6.0	1.5.2	1.5.0	1.4.1
Sqoop	-	-	-	-
Sqoop-Sandbox	-	-	-	-
TensorFlow	-	-	-	-

	emr-4.3.0	emr-4.2.0	emr-4.1.0	emr-4.0.0
Tez	-	-	-	-
Trino (PrestoSQL)	-	-	-	-
Zeppelin	-	-	-	-
Zeppelin-Sandbox	0.5.5	0.5.5	0.6.0-SNAPSHOT	-
ZooKeeper	-	-	-	-
ZooKeeper-Sandbox	-	-	-	-

4.0.0 릴리스 정보

4.0.0 구성 요소 버전

이 릴리스를 통해 Amazon EMR이 설치하는 구성 요소는 다음과 같습니다. 일부는 빅 데이터 애플리케이션 패키지의 일부로 설치됩니다. 나머지는 Amazon EMR에 고유하며 시스템 프로세스 및 기능을 위해 설치됩니다. 이는 일반적으로 `emr` 또는 `aws`로 시작됩니다. 최근 Amazon EMR 릴리스의 빅 데이터 애플리케이션 패키지는 일반적으로 커뮤니티에서 발견된 최신 버전입니다. 가능한 한 빨리 Amazon EMR에서 커뮤니티 릴리스를 제공합니다.

Amazon EMR의 일부 구성 요소는 커뮤니티 버전과 다릅니다. 이러한 구성 요소에는 *CommunityVersion*-amzn-*EmrVersion* 양식의 버전 레이블이 있습니다. *EmrVersion*은 0에서 시작합니다. 예를 들어, 버전 2.2의 `myapp-component`라는 오픈 소스 커뮤니티 구성 요소가 다른 Amazon EMR 릴리스에 포함되도록 세 번 수정된 경우 해당 릴리스 버전은 `2.2-amzn-2`로 나열됩니다.

구성 요소	버전	설명
emr-ddb	3.0.0	하둡 에코시스템 애플리케이션용 Amazon DynamoDB 커넥터

구성 요소	버전	설명
emr-goodies	2.0.0	편리한 하둡 에코시스템용 추가 라이브러리
emr-kinesis	3.0.0	하둡 에코시스템 애플리케이션용 Amazon Kinesis 커넥터.
emr-s3-dist-cp	2.0.0	Amazon S3용으로 최적화된 분사 복사 애플리케이션.
emrfs	2.0.0	하둡 에코시스템 애플리케이션용 Amazon S3 커넥터.
hadoop-client	2.6.0-amzn-0	'hdfs', 'hadoop', 'yarn' 등과 같은 하둡 명령줄 클라이언트
hadoop-hdfs-datanode	2.6.0-amzn-0	블록을 저장하는 HDFS 노드 수준 서비스
hadoop-hdfs-namenode	2.6.0-amzn-0	파일 이름 및 블록 위치를 추적하는 HDFS 서비스
hadoop-httpfs-server	2.6.0-amzn-0	HDFS 작업에 대한 HTTP 엔드포인트
hadoop-mapred	2.6.0-amzn-0	MapReduce 애플리케이션을 실행하는 MapReduce 실행 엔진 라이브러리
hadoop-yarn-nodemanager	2.6.0-amzn-0	개별 노드의 컨테이너를 관리하는 YARN 서비스
hadoop-yarn-resourcemanager	2.6.0-amzn-0	클러스터 리소스 및 분산 애플리케이션을 할당 및 관리하는 YARN 서비스
hive-client	1.0.0-amzn-0	Hive 명령줄 클라이언트

구성 요소	버전	설명
hive-metastore-server	1.0.0-amzn-0	하둡 작업에 대한 SQL용 메타데이터가 저장되는 시맨틱 리포지토리인 Hive 메타스토어에 액세스하기 위한 서비스
hive-server	1.0.0-amzn-0	Hive 쿼리를 웹 요청으로 수락하기 위한 서비스
mahout-client	0.10.0	머신 러닝을 위한 라이브러리.
mysql-server	5.5	MySQL 데이터베이스 서버.
pig-client	0.14.0-amzn-0	Pig 명령줄 클라이언트.
spark-client	1.4.1	Spark 명령줄 클라이언트.
spark-history-server	1.4.1	완료된 Spark 애플리케이션의 수명에 대해 기록된 이벤트를 조회하는 웹 UI.
spark-on-yarn	1.4.1	YARN용 인 메모리 실행 엔진.
spark-yarn-slave	1.4.1	YARN 슬레이브에서 필요한 Apache Spark 라이브러리.

4.0.0 구성 분류

구성 분류를 사용하면 애플리케이션을 사용자 지정할 수 있습니다. 이는 종종 `hive-site.xml`과 같이 애플리케이션의 구성 XML 파일에 해당합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

emr-4.0.0 분류

분류	설명
capacity-scheduler	하둡 capacity-scheduler.xml 파일의 값을 변경합니다.

분류	설명
core-site	하둡 core-site.xml 파일의 값을 변경합니다.
emrfs-site	EMRFS 설정을 변경합니다.
hadoop-env	모든 하둡 구성 요소에 대한 하둡 환경의 값을 변경합니다.
hadoop-log4j	하둡 log4j.properties 파일의 값을 변경합니다.
hdfs-site	HDFS hdfs-site.xml의 값을 변경합니다.
hive-env	Hive 환경의 값을 변경합니다.
hive-exec-log4j	Hive's hive-exec-log4j.properties 파일의 값을 변경합니다.
hive-log4j	Hive's hive-log4j.properties 파일의 값을 변경합니다.
hive-site	Hive hive-site.xml 파일의 값을 변경합니다.
httpfs-env	HTTPFS 환경의 값을 변경합니다.
httpfs-site	하둡 httpfs-site.xml 파일의 값을 변경합니다.
mapred-env	MapReduce 애플리케이션 환경의 값을 변경합니다.
mapred-site	MapReduce 애플리케이션 mapred-site.xml 파일의 값을 변경합니다.
pig-properties	Pig pig.properties 파일의 값을 변경합니다.
pig-log4j	Pig log4j.properties 파일의 값을 변경합니다.
spark	Apache Spark에 대한 Amazon EMR 큐레이트 설정입니다.

분류	설명
spark-defaults	Spark spark-defaults.conf 파일의 값을 변경합니다.
spark-env	the Spark 환경의 값을 변경합니다.
spark-log4j	Spark log4j.properties 파일의 값을 변경합니다.
yarn-env	YARN 환경의 값을 변경합니다.
yarn-site	YARN yarn-site.xml 파일의 값을 변경합니다.

Amazon EMR 2.x 및 3.x AMI 버전

> **Note**
>
> AWS 는 모든 AWS API 엔드포인트에 대한 TLS 구성을 TLS 1.2의 최소 버전으로 업데이트하고 있습니다. Amazon EMR 릴리스 3.10 이하 버전은 TLS 1.0 및 1.1 연결만 지원합니다. 2023년 12월 4일 이후 Amazon EMR 3.10 이하에서는 클러스터를 생성할 수 없습니다. Amazon EMR 3.10 이하를 사용하는 경우 워크로드를 즉시 테스트하고 최신 Amazon EMR 릴리스로 마이그레이션하는 것이 좋습니다. 자세한 내용은 AWS 보안 블로그를 참조하세요.

AMI 버전이라고 하는 Amazon EMR 2.x 및 3.x 릴리스는 호환성 때문에 해당 릴리스가 필요한 기존 솔루션에서 사용할 수 있습니다. 이 릴리스 버전으로 새 클러스터나 새 솔루션을 생성하는 것은 권장되지 않습니다. 여기에는 새 릴리스의 기능이 없으며 오래된 애플리케이션 패키지가 포함됩니다.

최신 Amazon EMR 릴리스 버전을 사용하여 솔루션을 구축하는 것이 좋습니다.

2.x 및 3.x 시리즈 릴리스 버전과 최신 Amazon EMR 릴리스 버전 간의 차이는 큽니다. 클러스터의 생성 및 구성 방법부터 클러스터의 애플리케이션 디렉터리 구조 및 포트에 이르기까지 매우 광범위합니다.

이 섹션에서는 Amazon EMR에 대한 가장 두드러진 차이와 특정 애플리케이션 구성 및 관리의 차이를 다룹니다. 이 내용은 포괄적이지 않습니다. 2.x 또는 3.x 시리즈에서 클러스터를 생성하여 사용하면 이 단원에서 설명하지 않은 차이가 생길 수 있습니다.

주제

Amazon EMR의 이전 AMI 버전을 포함하는 클러스터 생성

Amazon EMR 2.x 및 3.x 릴리스는 AMI 버전으로 참조됩니다. Amazon EMR 릴리스 4.0.0 이상에서 릴리스 레이블(예: `emr-5.11.0`)을 사용하여 릴리스 버전별로 릴리스를 참조합니다. 이 변경 사항은 AWS CLI 또는 프로그래밍 방식으로 클러스터를 생성할 때 가장 명확하게 나타납니다.

AWS CLI 를 사용하여 AMI 릴리스 버전을 사용하여 클러스터를 생성하는 경우 `--ami-version` 옵션을 사용합니다. 예: `--ami-version 3.11.0`. Amazon EMR 4.0.0 이상에 도입된 여러 옵션, 기능 및 애플리케이션은 `--ami-version`을 지정할 때 사용할 수 없습니다. 자세한 내용은 AWS CLI 명령 참조에서 create-cluster를 참조하세요.

다음 예제 AWS CLI 명령은 AMI 버전을 사용하여 클러스터를 시작합니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,\
InstanceType=m3.xlarge InstanceGroupType=CORE,InstanceCount=2,\
InstanceType=m3.xlarge --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-
actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

프로그래밍 방식에서는 모든 Amazon EMR 릴리스 버전이 EMR API에서 `RunJobFlowRequest` 작업을 사용하여 클러스터를 생성합니다. 다음은 AMI 릴리스 버전 3.11.0을 사용하여 클러스터를 생성하는 예제 Java 코드입니다.

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("AmiVersion Cluster")
   .withAmiVersion("3.11.0")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

다음 `RunJobFlowRequest` 호출은 릴리스 레이블을 대신 사용합니다.

```
RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("ReleaseLabel Cluster")
   .withReleaseLabel("emr-7.9.0")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myKeyPair")
    .withInstanceCount(1)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m3.xlarge")
    .withSlaveInstanceType("m3.xlarge");
```

클러스터 크기 구성

클러스터가 실행되면 Hadoop에서 데이터를 처리하는 데 필요한 매퍼 및 reducer 작업 수를 결정합니다. 클러스터의 규모가 클수록 리소스 사용과 처리 시간 단축을 위해 더 많은 작업을 수행해야 합니다. 일반적으로 EMR 클러스터는 전체 클러스터에서 동일한 크기를 유지합니다. 클러스터를 작성할 때 작업 수를 설정합니다. 실행 중인 클러스터의 크기를 조정할 때 클러스터 실행 중에 처리 방식을 변경할 수 있습니다. 따라서 고정된 수의 작업을 사용하는 대신 클러스터 수명 기간 동안 작업 수를 변경할 수 있습니다. 이상적인 작업 수를 설정하는 데 도움이 되는 두 가지 구성 옵션이 있습니다.

- `mapred.map.tasksperslot`
- `mapred.reduce.tasksperslot`

mapred-conf.xml 파일에서 두 옵션을 모두 설정할 수 있습니다. 작업을 클러스터에 제출하면 작업 클라이언트는 클러스터 전체에서 사용 가능한 현재 맵 및 reduce 슬롯의 총 수를 확인합니다. 작업 클라이언트는 다음 수식을 사용하여 작업 수를 설정합니다.

- mapred.map.tasks = mapred.map.tasksperslot * 클러스터의 맵 슬롯
- mapred.reduce.tasks =mapred.reduce.tasksperslot * 클러스터의 reduce 슬롯

작업 클라이언트는 작업 수가 구성되지 않은 경우 tasksperslot 파라미터만 읽습니다. 부트스트랩 작업을 통해 모든 클러스터에 대해 또는 구성을 변경하는 단계를 추가하여 작업마다 개별적으로 언제든지 작업 수를 재정의할 수 있습니다.

Amazon EMR은 태스크 노드 장애를 견디며 태스크 노드를 사용할 수 없게 되더라도 클러스터 실행을 계속합니다. Amazon EMR은 실패한 태스크 노드를 대체하기 위해 추가 태스크 노드를 자동으로 프로비저닝합니다.

클러스터 단계마다 작업 노드의 수가 다를 수 있습니다. 실행 중인 클러스터에 단계를 추가하여 작업 노드의 수를 수정할 수도 있습니다. 모든 단계는 기본적으로 순차적으로 실행되므로 어떤 단계에 대해서도 실행 중인 작업 노드 수를 지정할 수 있습니다.

Amazon EMR의 이전 AMI 버전을 사용하여 애플리케이션 설치

AMI 버전을 사용하는 경우 RunJobflow 작업의 NewSupportedProducts 파라미터 사용, 부트스트랩 작업 사용, Step 작업 사용 등 다양한 방법으로 애플리케이션이 설치됩니다.

Amazon EMR의 이전 AMI 버전을 사용하여 클러스터 및 애플리케이션 구성 사용자 지정

Amazon EMR 릴리스 버전 4.0.0에서는 구성 분류를 사용하여 애플리케이션을 구성하는 간단한 방법이 도입되었습니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오. AMI 버전을 사용할 때 전달하는 인수와 함께 부트스트랩 작업을 사용하여 애플리케이션을 구성합니다. 예를 들어, configure-hadoop 및 configure-daemons 부트스트랩 작업은 Hadoop 및 Yarn별 환경 속성(예: --namenode-heap-size)을 설정합니다. 최신 버전에서는 이러한 사항이 hadoop-env 및 yarn-env 구성 분류를 통해 구성됩니다. 일반적인 구성을 수행하는 부트스트랩 작업은 Github의 emr-bootstrap-actions 리포지토리를 참조하십시오.

다음 테이블에서는 부트스트랩 작업을 최신 Amazon EMR 릴리스 버전에 있는 구성 분류에 매핑합니다.

Hadoop

영향을 받는 애플리케이션 파일 이름	AMI 버전 부트스트랩 작업	구성 분류
core-site.xml	configure-hadoop -c	core-site
log4j.properties	configure-hadoop -l	hadoop-log4j
hdfs-site.xml	configure-hadoop -s	hdfs-site
해당 사항 없음	해당 사항 없음	hdfs-encryption-zones
mapred-site.xml	configure-hadoop -m	mapred-site
yarn-site.xml	configure-hadoop -y	yarn-site
httpfs-site.xml	configure-hadoop -t	httpfs-site
capacity-scheduler.xml	configure-hadoop -z	capacity-scheduler
yarn-env.sh	configure-daemons --resourcemanager-opts	yarn-env

Hive

영향을 받는 애플리케이션 파일 이름	AMI 버전 부트스트랩 작업	구성 분류
hive-env.sh	해당 사항 없음	hive-env
hive-site.xml	hive-script --install-hive-site ${MY_HIVE_SITE_FILE}	hive-site
hive-exec-log4j.properties	해당 사항 없음	hive-exec-log4j

영향을 받는 애플리케이션 파일 이름	AMI 버전 부트스트랩 작업	구성 분류
`hive-log4j.properties`	해당 사항 없음	`hive-log4j`

EMRFS

영향을 받는 애플리케이션 파일 이름	AMI 버전 부트스트랩 작업	구성 분류
`emrfs-site.xml`	`configure-hadoop -e`	`emrfs-site`
해당 사항 없음	`s3get -s s3://custom-provider.jar -d /usr/share/aws/emr/auxlib/`	`emrfs-site` (새로운 설정 `fs.s3.cse.encryptionMaterialsProvider.uri` 지원)

모든 분류 목록을 보려면 [애플리케이션 구성](#) 단원을 참조하십시오.

애플리케이션 환경 변수

AMI 버전을 사용할 때 하둡 환경을 구성하기 위해 `hadoop-user-env.sh` 스크립트가 `configure-daemons` 부트스트랩 작업과 함께 사용됩니다. 이 스크립트에는 다음 작업이 포함되어 있습니다.

```
#!/bin/bash
export HADOOP_USER_CLASSPATH_FIRST=true;
echo "HADOOP_CLASSPATH=/path/to/my.jar" >> /home/hadoop/conf/hadoop-user-env.sh
```

Amazon EMR 릴리스 4.x에서는 다음 예제에 표시된 대로 `hadoop-env` 구성 분류를 사용하여 동일하게 수행할 수 있습니다.

```
[
      {
         "Classification":"hadoop-env",
         "Properties":{

         },
```

```
            "Configurations":[
               {
                  "Classification":"export",
                  "Properties":{
                     "HADOOP_USER_CLASSPATH_FIRST":"true",
                     "HADOOP_CLASSPATH":"/path/to/my.jar"
                  }
               }
            ]
         }
      ]
```

또 다른 예로, `configure-daemons`를 사용하고 `--namenode-heap-size=2048` 및 `--namenode-opts=-XX:GCTimeRatio=19`를 전달하는 것은 다음 구성 분류와 동일합니다.

```
[
      {
         "Classification":"hadoop-env",
         "Properties":{

         },
         "Configurations":[
            {
               "Classification":"export",
               "Properties":{
                  "HADOOP_DATANODE_HEAPSIZE":  "2048",
            "HADOOP_NAMENODE_OPTS":  "-XX:GCTimeRatio=19"
               }
            }
         ]
      }
   ]
```

다른 애플리케이션 환경 변수는 더 이상 `/home/hadoop/.bashrc`에 정의되지 않습니다. 대신 `/etc/default`과 같이 구성 요소별 또는 애플리케이션별 `/etc/default/hadoop` 파일에서 주로 설정됩니다. 또한 애플리케이션 RPM에서 설치한 `/usr/bin/`의 래퍼 스크립트는 실제 빈 스크립트를 포함하기 전에 추가 환경 변수를 설정할 수 있습니다.

서비스 포트

AMI 버전을 사용할 때 일부 서비스에서는 사용자 지정 포트를 사용합니다.

포트 설정의 변경

설정	AMI 버전 3.x	오픈 소스 기본값
fs.default.name	hdfs://emrDeterminedIP:9000	기본값(hdfs://*emrDeterminedIP* :8020)
dfs.datanode.address	0.0.0.0:9200	기본값(0.0.0.0:50010)
dfs.datanode.http.address	0.0.0.0:9102	기본값(0.0.0.0:50075)
dfs.datanode.https.address	0.0.0.0:9402	기본값(0.0.0.0:50475)
dfs.datanode.ipc.address	0.0.0.0:9201	기본값(0.0.0.0:50020)
dfs.http.address	0.0.0.0:9101	기본값(0.0.0.0:50070)
dfs.https.address	0.0.0.0:9202	기본값(0.0.0.0:50470)
dfs.secondary.http.address	0.0.0.0:9104	기본값(0.0.0.0:50090)
yarn.nodemanager.address	0.0.0.0:9103	기본값(${yarn.nodemanager.hostname}:0)
yarn.nodemanager.localizer.address	0.0.0.0:9033	기본값(${yarn.nodemanager.hostname}:8040)
yarn.nodemanager.webapp.address	0.0.0.0:9035	기본값(${yarn.nodemanager.hostname}:8042)
yarn.resourcemanager.address	*emrDeterminedIP* :9022	기본값(${yarn.resourcemanager.hostname}:8032)
yarn.resourcemanager.admin.address	*emrDeterminedIP* :9025	기본값(${yarn.resourcemanager.hostname}:8033)
yarn.resourcemanager.resource-tracker.address	*emrDeterminedIP* :9023	기본값(${yarn.resourcemanager.hostname}:8031)
yarn.resourcemanager.scheduler.address	*emrDeterminedIP* :9024	기본값(${yarn.resourcemanager.hostname}:8030)

설정	AMI 버전 3.x	오픈 소스 기본값
yarn.resourcemanager.webapp.address	0.0.0.0:9026	기본값(${yarn.resourcemanager.hostname}:8088)
yarn.web-proxy.address	*emrDeterminedIP* :9046	기본값(값 지정 안 함)
yarn.resourcemanager.hostname	0.0.0.0(기본값)	*emrDeterminedIP*

> **Note**
>
> *emrDeterminedIP*는 Amazon EMR에서 생성되는 IP 주소입니다.

Users

AMI 버전을 사용할 때 hadoop 사용자는 모든 프로세스를 실행하고 모든 파일을 소유합니다. Amazon EMR 릴리스 버전 4.0.0 이상에서는 사용자가 애플리케이션 및 구성 요소 수준에 존재합니다.

설치 시퀀스, 설치된 아티팩트 및 로그 파일 위치

AMI 버전을 사용할 때 애플리케이션 아티팩트 및 구성 요소 디렉터리는 /home/hadoop/*application* 디렉터리에 설치됩니다. 예를 들어 Hive를 설치한 경우 디렉터리는 /home/hadoop/hive입니다. Amazon EMR 릴리스 4.0.0 이상에서는 애플리케이션 아티팩트가 /usr/lib/*application* 디렉터리에 설치됩니다. AMI 버전을 사용할 때 로그 파일은 다양한 위치에 저장됩니다. 아래 표에는 위치가 나열되어 있습니다.

Amazon S3에서 로그 위치의 변경

대몬(daemon) 또는 애플리케이션	디렉터리 위치
인스턴스 상태	node/*instance-id* /instance-state/
hadoop-hdfs-namenode	daemons/*instance-id* /hadoop-hadoop-namenode.log
hadoop-hdfs-datanode	daemons/*instance-id* /hadoop-hadoop-datanode.log

대몬(daemon) 또는 애플리케이션	디렉터리 위치
hadoop-yarn(ResourceManager)	daemons/*instance-id* /yarn-hadoop-resourcemanager
하둡 yarn(프록시 서버)	daemons/*instance-id* /yarn-hadoop-proxyserver
mapred-historyserver	daemons/*instance-id* /
httpfs	daemons/*instance-id* /httpfs.log
hive-server	node/*instance-id* /hive-server/hive-server.log
hive-metastore	node/*instance-id* /apps/hive.log
Hive CLI	node/*instance-id* /apps/hive.log
YARN 애플리케이션 사용자 로그 및 컨테이너 로그	task-attempts/
Mahout	N/A
Pig	N/A
spark-historyserver	해당 사항 없음
mapreduce 작업 기록 파일	jobs/

Command Runner

AMI 버전을 사용할 때 `/home/hadoop/contrib/streaming/hadoop-streaming.jar` 등의 여러 스크립트 또는 프로그램이 셸 로그인 경로 환경에 배치되지 않으므로, command-runner.jar 또는 script-runner.jar 등 jar 파일을 사용하여 스크립트를 실행할 경우 전체 경로를 지정해야 합니다. `command-runner.jar`은 AMI에 있으므로 `script-runner.jar`의 경우처럼 전체 URI를 알 필요가 없습니다.

복제 인수

복제 인수를 사용하면 하둡 JVM을 시작할 시기를 구성할 수 있습니다. 모든 작업에 대해 새 하둡 JVM을 시작하여 작업을 보다 효과적으로 격리하거나, 작업 간에 JVM을 공유하여 프레임워크 오버헤드를 낮출 수 있습니다. 많은 작은 파일을 처리하려는 경우에는 JVM을 여러 번 재사용하여 시작 비용을 분할 상환하는 것이 합리적입니다. 하지만 각 작업마다 오랜 시간이 걸리거나 대량의 데이터를 처리하는 경우에는 모든 메모리가 후속 작업을 위해 해제될 수 있도록 JVM을 재사용하지 않도록 선택할 수 있습니다. AMI 버전을 사용할 때 `configure-hadoop` 부트스트랩 작업을 사용하여 복제 인수를 `mapred.job.reuse.jvm.num.tasks` 속성으로 설정하도록 사용자 지정할 수 있습니다.

다음은 무제한 JVM 재사용을 위해 JVM 재사용 인수를 설정한 예입니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.11.0 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m3.xlarge \
InstanceGroupType=CORE,InstanceCount=2,InstanceType=m3.xlarge \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hadoop,\
Name="Configuring infinite JVM reuse",Args=["-m","mapred.job.reuse.jvm.num.tasks=-1"]
```

Amazon EMR의 이전 AMI 버전에 고유한 Hive 애플리케이션

로그 파일

Amazon EMR AMI 버전 2.x 및 3.x를 사용하면 Hive 로그가 `/mnt/var/log/apps/`에 저장됩니다. Hive의 동시 버전을 지원하려면 다음 표와 같이 실행하는 Hive 버전에 따라 로그 파일 이름이 결정됩니다.

Hive 버전	로그 파일 이름
0.13.1	`hive.log`

Hive 버전	로그 파일 이름
	Note 이 버전부터 Amazon EMR은 버전 관리 미사용 파일 이름(`hive.log`)을 사용합니다. 마이너 버전은 메이저 버전과 동일한 로그 위치를 공유합니다.
0.11.0	`hive_0110.log` **Note** Hive 0.11.0의 마이너 버전(예: 0.11.0.1)은 Hive 0.11.0과 동일한 로그 파일 위치를 공유합니다.
0.8.1	`hive_081.log` **Note** Hive 0.8.1.1의 마이너 버전(예: 0.8.1.1)은 Hive 0.8.1과 동일한 로그 파일 위치를 공유합니다.
0.7.1	`hive_07_1.log` **Note** Hive 0.7.1의 마이너 버전(예: Hive 0.7.1.3 및 Hive 0.7.1.4)은 Hive 0.7.1과 동일한 로그 파일 위치를 공유합니다.
0.7	`hive_07.log`
0.5	`hive_05.log`
0.4	`hive.log`

분할 입력 기능

0.13.1 이전의 Hive 버전(3.11.0 이전의 Amazon EMR AMI 버전)을 사용하여 분할 입력 기능을 구현하려면 다음을 사용합니다.

```
hive> set hive.input.format=org.apache.hadoop.hive.ql.io.HiveCombineSplitsInputFormat;
hive> set mapred.min.split.size=100000000;
```

이 기능은 Hive 0.13.1에서 더 이상 사용되지 않습니다. Amazon EMR AMI 버전 3.11.0에서 동일한 분할 입력 형식 기능을 사용하려면 다음을 사용합니다.

```
set hive.hadoop.supports.splittable.combineinputformat=true;
```

Thrift 서비스 포트

Thrift는 나중에 분석할 수 있도록 데이터 구조를 유지하는 데 사용되는 압축된 이진 직렬화 형식을 정의하는 RPC 프레임워크입니다. 일반적으로 Hive는 다음 포트에서 작동하도록 서버를 구성합니다.

Hive 버전	포트 번호
Hive 0.13.1	10000
Hive 0.11.0	10004
Hive 0.8.1	10003
Hive 0.7.1	10002
Hive 0.7	10001
Hive 0.5	10000

Thrift 서비스에 대한 자세한 내용은 http://wiki.apache.org/thrift/를 참조하십시오.

Hive를 사용하여 파티션 복구

Amazon EMR에는 Amazon S3에 있는 테이블 데이터에서 테이블의 파티션을 복구하는 Hive 쿼리 언어의 문이 포함되어 있습니다. 다음 예제는 이 작업을 보여 줍니다.

```
CREATE EXTERNAL TABLE (json string) raw_impression
PARTITIONED BY (dt string)
LOCATION 's3://elastic-mapreduce/samples/hive-ads/tables/impressions';
ALTER TABLE logs RECOVER PARTITIONS;
```

파티션 디렉터리 및 데이터가 테이블 정의에 지정된 위치에 있어야 하며 Hive 규칙에 따라 이름이 지정되어야 합니다. 예: `dt=2009-01-01`.

> **Note**
>
> Hive 0.13.1 이후에는 이 기능이 `msck repair table`을 사용하여 기본적으로 지원되므로 `recover partitions`가 지원되지 않습니다. 자세한 내용은 https://cwiki.apache.org/confluence/display/Hive/LanguageManual+DDL을 참조하십시오.

스크립트에 Hive 변수 전달

를 사용하여 Hive 단계에 변수를 전달하려면 다음 명령을 AWS CLI입력하고 *myKey*를 EC2 키 페어의 이름으로 바꾸고 *amzn-s3-demo-bucket*을 버킷 이름으로 바꿉니다. 이 예에서 `SAMPLE`은 앞에 `-d` 스위치가 붙는 변수 값입니다. 이 변수는 Hive 스크립트에서 `${SAMPLE}`로 정의됩니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.9 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,\
Args=[-f,s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,\
INPUT=s3://elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/
hive-ads/output/,\
-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

외부 메타스토어 위치 지정

다음 절차에서는 Hive 메타스토어 위치에 대한 기본 구성 값을 재정의하고 재구성된 메타스토어 위치를 사용하여 클러스터를 시작하는 방법을 보여 줍니다.

EMR 클러스터 외부에 위치한 메타스토어를 생성하려면

1. Amazon RDS를 사용하여 MySQL 또는 Aurora 데이터베이스를 생성합니다.

   Amazon RDS 데이터베이스 생성 방법에 대한 자세한 내용은 [Amazon RDS 시작하기]를 참조하세요.

2. 데이터베이스와 ElasticMapReduce-Master 보안 그룹 간의 JDBC 연결을 허용하도록 보안 그룹을 수정합니다.

   액세스를 위해 보안 그룹을 수정하는 방법에 대한 자세한 내용은 Amazon RDS 사용 설명서에서 [Amazon RDS 보안 그룹]을 참조하세요.

3. `hive-site.xml`에서 JDBC 구성 값을 설정합니다.

   a. 다음 사항이 포함된 `hive-site.xml` 구성 파일을 생성합니다.

```
<configuration>
  <property>
    <name>javax.jdo.option.ConnectionURL</name>
    <value>jdbc:mariadb://hostname:3306/hive?createDatabaseIfNotExist=true</
value>
    <description>JDBC connect string for a JDBC metastore</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionUserName</name>
    <value>hive</value>
    <description>Username to use against metastore database</description>
  </property>
  <property>
    <name>javax.jdo.option.ConnectionPassword</name>
    <value>password</value>
    <description>Password to use against metastore database</description>
  </property>
</configuration>
```

   *`hostname`*은 데이터베이스를 실행하는 Amazon RDS 인스턴스의 DNS 주소입니다. *`username`* 및 *`password`*는 데이터베이스의 보안 인증입니다. MySQL 및 Aurora 데이터베이

스 인스턴스에 연결하는 방법에 대한 자세한 내용은 Amazon RDS 사용 설명서에서 MySQL 데이터베이스 엔진 기반 DB 인스턴스에 연결하기 및 Aurora DB 클러스터에 연결을 참조하세요.

JDBC 드라이버는 Amazon EMR에서 설치됩니다.

> **Note**
>
> 값 속성에는 공백이나 캐리지 리턴이 포함되어서는 안 됩니다. 속성은 모두 한 줄에 나타나야 합니다.

b. `hive-site.xml` 파일을 Amazon S3 위의 한 위치로 이동합니다(예: `s3://amzn-s3-demo-bucket/hive-site.xml`).

4. 클러스터를 생성하고 사용자 지정된 `hive-site.xml` 파일의 Amazon S3 위치를 지정합니다.

   다음 예제 명령은이 작업을 수행하는 AWS CLI 명령을 보여줍니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.10 \
--applications Name=Hue Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 \
--bootstrap-actions Name="Install Hive Site Configuration",\
Path="s3://region.elasticmapreduce/libs/hive/hive-script",\
Args=["--base-path","s3://elasticmapreduce/libs/hive","--install-hive-site",\
"--hive-site=s3://amzn-s3-demo-bucket/hive-site.xml","--hive-versions","latest"]
```

JDBC를 사용하여 Hive에 연결

JDBC를 통해 Hive에 연결하려면 JDBC 드라이버를 다운로드하고 SQL 클라이언트를 설치해야 합니다. 다음 예제에서는 SQL Workbench/J에서 JDBC를 사용하여 Hive에 연결하는 방법을 보여 줍니다.

JDBC 드라이버를 다운로드하려면

1. 액세스하려는 Hive 버전에 적합한 드라이버를 다운로드하고 압축을 풉니다. Hive 버전은 Amazon EMR 클러스터를 생성할 때 선택하는 AMI에 따라 다릅니다.
   - Hive 0.13.1 JDBC 드라이버: https://amazon-odbc-jdbc-drivers.s3.amazonaws.com/public/AmazonHiveJDBC_1.0.4.1004.zip
   - Hive 0.11.0 JDBC 드라이버: https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.11.0
   - Hive 0.8.1 JDBC 드라이버: https://mvnrepository.com/artifact/org.apache.hive/hive-jdbc/0.8.1
2. 자세한 내용은 SQL Workbench/J 사용 설명서에서 Installing and starting SQL Workbench/J를 참조하십시오.
3. 클러스터 마스터 노드에 대한 SSH 터널을 생성합니다. 연결용 포트는 Hive 버전에 따라 달라집니다. 예제 명령은 Linux ssh 사용자의 경우 아래 표에, Windows 사용자의 경우 PuTTY 명령에 제공됩니다.

   Linux SSH 명령

Hive 버전	Command
0.13.1	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000 hadoop@ master-public-dns-name`
0.11.0	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10004:localhost:10004 hadoop@ master-public-dns-name`
0.8.1	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10003:localhost:10003 hadoop@ master-public-dns-name`
0.7.1	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10002:localhost:10002 hadoop@ master-public-dns-name`
0.7	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10001:localhost:10001 hadoop@ master-public-dns-name`
0.5	`ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000 hadoop@ master-public-dns-name`

Windows PuTTY 터널 설정

Hive 버전	터널 설정
0.13.1	소스 포트: 10000 대상: *master-public-dns-name* :10000
0.11.0	소스 포트: 10004 대상: *master-public-dns-name* :10004
0.8.1	소스 포트: 10003 대상: *master-public-dns-name* :10003

4. JDBC 드라이버를 SQL Workbench에 추가합니다.

   a. Select Connection Profile(연결 프로필 선택) 대화 상자에서 Manage Drivers(드라이버 관리)를 선택합니다.

   b. Create a new entry(새 항목 생성)(빈 페이지) 아이콘을 선택합니다.

   c. Name(이름) 필드에 **Hive JDBC**을(를) 입력합니다.

   d. 라이브러리에서 Select the JAR file(s)(JAR 파일 선택) 아이콘을 클릭합니다.

   e. 다음 표에 표시된 대로 JAR 파일을 선택합니다.

Hive 드라이버 버전	추가할 JAR 파일

0.13.1	`hive_metastore.jar`
	`hive_service.jar`
	`HiveJDBC3.jar`
	`libfb303-0.9.0.jar`
	`libthrift-0.9.0.jar`
	`log4j-1.2.14.jar`
	`ql.jar`
	`slf4j-api-1.5.8.jar`
	`slf4j-log4j12-1.5.8.jar`
	`TCLIServiceClient.jar`

0.11.0	`hadoop-core-1.0.3.jar`
	`hive-exec-0.11.0.jar`
	`hive-jdbc-0.11.0.jar`
	`hive-metastore-0.11.0.jar`
	`hive-service-0.11.0.jar`

Hive 드라이버 버전	추가할 JAR 파일
	libfb303-0.9.0.jar commons-logging-1.0.4.jar slf4j-api-1.6.1.jar
0.8.1	hadoop-core-0.20.205.jar hive-exec-0.8.1.jar hive-jdbc-0.8.1.jar hive-metastore-0.8.1.jar hive-service-0.8.1.jar libfb303-0.7.0.jar libthrift-0.7.0.jar log4j-1.2.15.jar slf4j-api-1.6.1.jar slf4j-log4j12-1.6.1.jar
0.7.1	hadoop-0.20-core.jar hive-exec-0.7.1.jar hive-jdbc-0.7.1.jar hive-metastore-0.7.1.jar hive-service-0.7.1.jar libfb303.jar commons-logging-1.0.4.jar slf4j-api-1.6.1.jar slf4j-log4j12-1.6.1.jar
0.7	hadoop-0.20-core.jar hive-exec-0.7.0.jar hive-jdbc-0.7.0.jar hive-metastore-0.7.0.jar hive-service-0.7.0.jar libfb303.jar commons-logging-1.0.4.jar slf4j-api-1.5.6.jar slf4j-log4j12-1.5.6.jar

Hive 드라이버 버전	추가할 JAR 파일

0.5	`hadoop-0.20-core.jar`
	`hive-exec-0.5.0.jar`
	`hive-jdbc-0.5.0.jar`
	`hive-metastore-0.5.0.jar`
	`hive-service-0.5.0.jar`
	`libfb303.jar`
	`log4j-1.2.15.jar`
	`commons-logging-1.0.4.jar`

f. Please select one driver(드라이버 하나를 선택하십시오) 대화 상자에서 다음 표에 해당하는 드라이버를 선택하고 확인을 클릭합니다.

Hive 버전	드라이버 클래스 이름
0.13.1	`com.amazon.hive.jdbc3.HS2Driver`
0.11.0	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.8.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7.1	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.7	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`
0.5	`org.apache.hadoop.hive.jdbc.HiveDriver.jar`

5. Select Connection Profile(연결 프로파일 선택) 대화 상자로 돌아가면 드라이버 필드가 Hive JDBC로 설정되어 있는지 확인하고 다음 표에 해당하는 URL 필드에 JDBC 연결 문자열을 제공합니다.

Hive 버전	JDBC 연결 문자열
0.13.1	`jdbc:hive2://localhost:10000/default`
0.11.0	`jdbc:hive://localhost:10004/default`
0.8.1	`jdbc:hive://localhost:10003/default`

클러스터가 AMI 버전 3.3.1 이상을 사용하는 경우 연결 프로파일 선택 대화 상자에서 사용자 이름 필드에 **hadoop**을 입력합니다.

Amazon EMR의 이전 AMI 버전에 고유한 HBase 애플리케이션

지원되는 HBase 버전

HBase 버전	AMI 버전	AWS CLI 구성 파라미터	HBase 버전 세부 정보

0.94.18	3.1.0 이상	`--ami-version 3.1`
		`--ami-version 3.2`
		`--ami-version 3.3`
		`--applications Name=HBase`
			• 버그 수정 사항 및 기능 향상.

0.94.7	3.0-3.0.4	`--ami-version 3.0` `--applications Name=HBase`	

HBase 버전	AMI 버전	AWS CLI 구성 파라미터	HBase 버전 세부 정보

0.92	2.2 이상	`--ami-version 2.2 or later`
		
		`--applications Name=HBase`	

HBase 클러스터 필수 조건

Amazon EMR AMI 버전 2.x 및 3.x를 사용하여 생성된 클러스터는 HBase에 대한 다음 요구 사항을 충족해야 합니다.

- AWS CLI (선택 사항) - 명령줄을 사용하여 HBase와 상호 작용하려면 최신 버전의를 다운로드하여 설치합니다 AWS CLI. 자세한 내용을 알아보려면 AWS Command Line Interface 사용자 가이드에서 AWS Command Line Interface설치를 참조하세요.
- 최소 두 개의 인스턴스(선택 사항) - 클러스터의 프라이머리 노드는 HBase 마스터 서버와 Zookeeper를 실행하고, 태스크 노드는 HBase 리전 서버를 실행합니다. 최상의 성능을 위해 HBase 클러스터는 최소 두 개의 EC2 인스턴스에서 실행해야 하지만 평가를 위해 단일 노드에서 HBase를 실행할 수 있습니다.
- 장기 실행 클러스터 - HBase는 장기 실행 클러스터에서만 실행됩니다. 기본적으로 CLI 및 Amazon EMR 콘솔은 장기 실행 클러스터를 생성합니다.
- Amazon EC2 키 페어 세트(권장) - Secure Shell(SSH) 네트워크 프로토콜을 사용하여 프라이머리 노드에 연결하고 HBase 쉘 명령을 실행하려면 클러스터를 생성할 때 Amazon EC2 키 페어를 사용해야 합니다.
- 올바른 AMI 및 Hadoop 버전 - HBase 클러스터는 현재 Hadoop 20.205 이상에서만 지원됩니다.
- Ganglia(선택 사항) - HBase 성능 지표를 모니터링하려면 클러스터를 생성할 때 Ganglia를 설치합니다.
- 로그용 Amazon S3 버킷(선택 사항) - HBase의 로그는 프라이머리 노드에서 사용할 수 있습니다. 이 로그를 Amazon S3에 복사하려면 클러스터를 생성할 때 로그 파일을 수신하는 S3 버킷을 지정합니다.

HBase를 포함하는 클러스터 생성

다음 테이블에는 콘솔에서 Amazon EMR AMI 릴리스 버전을 사용하여 HBase가 설치된 클러스터를 생성할 때 사용할 수 있는 옵션이 나열되어 있습니다.

필드	작업
백업에서 복원	Amazon S3에 저장된 데이터를 포함하는 HBase 클러스터를 미리 로드할지 여부를 지정합니다.
백업 위치	복원할 백업이 있는 Amazon S3의 URI를 지정합니다.
Backup version(백업 버전)	선택적으로 백업 위치에서 사용할 백업의 버전 이름을 지정합니다. 이 필드를 비워두면 Amazon EMR에서 백업 위치의 최신 백업을 사용하여 새 HBase 클러스터를 채웁니다.
Schedule Regular Backups(정기 백업 예약)	자동 증분 백업을 예약할지 여부를 지정합니다. 첫 번째 백업은 이후의 증분 백업에 대한 기준을 생성하기 위한 전체 백업입니다.
Consistent backup(일관적 백업)	백업의 일관성 여부를 지정합니다. 일관된 백업은 초기 백업 단계, 즉 노드 간 동기화 중에 쓰기 작업을 일시 중지하는 백업입니다. 이렇게 일시 중지된 쓰기 작업은 대기열에 배치되고 동기화가 완료되면 다시 시작됩니다.
Backup frequency(백업 빈도)	예약된 백업 간의 일/시간/분입니다.
백업 위치	백업이 저장되는 Amazon S3 URI입니다. 올바른 차등 백업을 유지하려면 각 HBase 클러스터의 백업 위치가 달라야 합니다.
Backup start time(백업 시작 시간)	첫 번째 백업의 발생 시점을 지정합니다. 클러스터가 실행되는 즉시 첫 번째 백업이 시작되도록 now로 설정하거나, ISO 형식으로 날짜와 시간을 입력할 수 있습니다. 예를 들어, 2012-06-15T20:00Z는 시작 시간을 2012년 6월 15일 오후 8시(UTC)로 설정합니다.

다음 예제 AWS CLI 명령은 HBase 및 기타 애플리케이션을 사용하여 클러스터를 시작합니다.

Note

가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
               --applications Name=Hue Name=Hive Name=Pig Name=HBase \
               --use-default-roles --ec2-attributes KeyName=myKey \
               --instance-type c1.xlarge --instance-count 3 --termination-protected
```

이전 절차에서처럼 Hive 및 HBase 클러스터 간의 연결이 설정된 후 Hive에 외부 테이블을 만들어서 HBase 클러스터에 저장된 데이터에 액세스할 수 있습니다.

Hive 프롬프트에서 실행할 때 다음 예제는 `inputTable`이라는 HBase 테이블에 저장된 데이터를 참조하는 외부 테이블을 생성합니다. 그러면 Hive 명령문의 `inputTable`을 참조하여 HBase 클러스터에 저장된 데이터를 쿼리 및 수정할 수 있습니다.

Note

다음 예제에서는 AMI 2.3.3의 protobuf-java-2.4.0a.jar을 사용하지만, 버전에 일치하도록 예제를 수정해야 합니다. 보유하고 있는 프로토콜 버퍼 JAR의 버전을 확인하려면 Hive 명령 프롬프트에서 `! ls /home/hadoop/lib;` 명령을 실행합니다.

```
add jar lib/emr-metrics-1.0.jar ;
               add jar lib/protobuf-java-2.4.0a.jar ;

               set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com ;

               create external table inputTable (key string, value string)
                    stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
                     with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
                     tblproperties ("hbase.table.name" = "t1");

               select count(*) from inputTable ;
```

HBase 구성 사용자 지정

대부분의 애플리케이션에서는 기본 설정이 작동해야 하지만 HBase 구성 설정을 유연하게 수정할 수 있습니다. 이렇게 하려면 다음 두 개의 부트스트랩 작업 스크립트 중 하나를 실행합니다.

- configure-hbase-daemons - master, regionserver, zookeeper daemons의 속성을 구성합니다. 이러한 속성에는 HBase 데몬이 시작될 때 JVM(Java Virtual Machine)에 전달할 힙 크기와 옵션이 포함됩니다. 부트스트랩 작업에서 이러한 속성을 인수로 설정합니다. 이 부트스트랩 작업은 HBase 클러스터의 /home/hadoop/conf/hbase-user-env.sh 구성 파일을 수정합니다.
- configure-hbase - HBase 마스터가 바인딩해야 하는 포트, 클라이언트 CLI가 작업을 재시도해야 하는 최대 횟수 등 HBase 사이트별 설정을 구성합니다. 이를 부트스트랩 작업의 인수로 하나씩 설정하거나 Amazon S3에서 XML 구성 파일의 위치를 지정할 수 있습니다. 이 부트스트랩 작업은 HBase 클러스터의 /home/hadoop/conf/hbase-site.xml 구성 파일을 수정합니다.

> **Note**
>
> 이러한 스크립트는 다른 부트스트랩 작업과 마찬가지로 클러스터를 만들 때만 실행할 수 있습니다. 현재 실행중인 HBase 클러스터의 구성을 변경하는 데 사용할 수 없습니다.

configure-hbase 또는 configure-hbase-daemons 부트스트랩 작업을 실행할 경우 지정한 값이 기본값을 재정의합니다. 명시적으로 설정하지 않은 값은 기본값을 받습니다.

이러한 부트스트랩 작업으로 HBase를 구성하는 것은 Amazon EMR에서 부트스트랩 작업을 사용하여 Hadoop 설정 및 Hadoop 대몬(daemon) 속성을 구성하는 것과 유사합니다. 차이점은 HBase에는 프로세스별 메모리 옵션이 없다는 것입니다. 대신 `--daemon-opts` 인수를 사용하여 메모리 옵션을 설정합니다. 여기서 *daemon*은 구성할 데몬의 이름으로 바뀝니다.

HBase 대몬(daemon) 구성

Amazon EMR은 HBase 대몬(daemon)의 구성을 변경하는 데 사용할 수 있는 부트스트랩 작업 `s3://region.elasticmapreduce/bootstrap-actions/configure-hbase-daemons`를 제공합니다. 여기서 *region*은 HBase 클러스터가 시작되는 리전입니다.

를 사용하여 HBase 데몬을 구성하려면 클러스터를 시작할 때 `configure-hbase-daemons` 부트스트랩 작업을 AWS CLI추가하여 하나 이상의 HBase 데몬을 구성합니다. 다음 속성을 설정할 수 있습니다.

속성	설명
`hbase-master-opts`	JVM이 마스터 데몬을 실행하는 방식을 제어하는 옵션입니다. 설정된 경우 기본 HBASE_MASTER_OPTS 변수를 재정의합니다.
`regionserver-opts`	JVM이 리전 서버 데몬을 실행하는 방식을 제어하는 옵션입니다. 설정된 경우 기본 HBASE_REGIONSERVER_OPTS 변수를 재정의합니다.
`zookeeper-opts`	JVM이 zookeeper 데몬을 실행하는 방식을 제어하는 옵션입니다. 설정된 경우 기본 HBASE_ZOOKEEPER_OPTS 변수를 재정의합니다.

이 옵션에 대한 자세한 내용은 HBase 설명서의 hbase-env.sh를 참조하십시오.

`zookeeper-opts` 및 `hbase-master-opts`에 대한 값을 구성하기 위한 부트스트랩 작업은 다음 예에 표시되어 있습니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-daemons,\
Args=["--hbase-zookeeper-opts=-Xmx1024m -XX:GCTimeRatio=19","--hbase-master-opts=-Xmx2048m","--hbase-regionserver-opts=-Xmx4096m"]
```

HBase 사이트 설정 구성

Amazon EMR은 HBase의 구성을 변경하는 데 사용할 수 있는 부트스트랩 작업 `s3://elasticmapreduce/bootstrap-actions/configure-hbase`를 제공합니다. 구성 값을 부트스

트랩 작업의 인수로 하나씩 설정하거나 Amazon S3에서 XML 구성 파일의 위치를 지정할 수 있습니다. 일부 구성 설정만 설정해야 하는 경우 구성 값을 하나씩 설정하는 것이 유용합니다. 변경해야 할 사항이 많거나 재사용할 수 있도록 구성 설정을 저장하려는 경우 XML 파일을 사용하여 설정하는 것이 유용합니다.

> **Note**
>
> Amazon S3 버킷 이름의 접두사로 리전 접두사를 붙일 수 있습니다(예: `s3://region.elasticmapreduce/bootstrap-actions/configure-hbase`) 여기서 *region*은 HBase 클러스터를 시작하는 리전입니다.

이 부트스트랩 작업은 HBase 클러스터의 `/home/hadoop/conf/hbase-site.xml` 구성 파일을 수정합니다. HBase 클러스터가 시작될 경우에만 부트스트랩 작업을 실행할 수 있습니다.

구성할 수 있는 HBase 사이트 설정에 대한 자세한 내용은 HBase 설명서에서 Default configuration을 참조하세요.

HBase 클러스터 시작 시 `configure-hbase` 부트스트랩 작업을 설정하고 변경할 `hbase-site.xml` 내에서 값을 지정합니다.

를 사용하여 개별 HBase 사이트 설정을 지정하려면 AWS CLI

- `hbase.hregion.max.filesize` 설정을 변경하려면 다음 명령을 입력하고 *myKey*를 Amazon EC2 키 페어의 이름으로 바꿉니다.

  > **Note**
  >
  > 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

  ```
  aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
  --applications Name=Hue Name=Hive Name=Pig Name=HBase \
  --use-default-roles --ec2-attributes KeyName=myKey \
  --instance-type c1.xlarge --instance-count 3 --termination-protected \
  --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
  hbase,Args=["-s","hbase.hregion.max.filesize=52428800"]
  ```

를 사용하여 XML 파일로 HBase 사이트 설정을 지정하려면 AWS CLI

1. 사용자 지정 `hbase-site.xml` 버전을 생성합니다. 사용자 지정 파일은 유효한 XML이어야 합니다. 오류 발생 가능성을 줄이려면 파일을 처음부터 새로 생성하는 대신에 `/home/hadoop/conf/hbase-site.xml`의 Amazon EMR HBase 프라이머리 노드에 위치한 `hbase-site.xml`의 기본 복사본으로 시작하여 해당 파일의 복사본을 편집합니다. 새 파일에 새 이름을 지정하거나 `hbase-site.xml`를 그대로 사용할 수 있습니다.
2. 사용자 지정 `hbase-site.xml` 파일을 Amazon S3 버킷에 업로드합니다. 클러스터를 시작하는 AWS 계정이 파일에 액세스할 수 있도록 권한이 설정되어 있어야 합니다. 클러스터를 시작하는 AWS 계정이 Amazon S3 버킷도 소유한 경우 액세스할 수 있습니다.
3. HBase 클러스터 시작 시 configure-hbase 부트스트랩 작업을 설정하고 사용자 지정 `hbase-site.xml` 파일의 위치를 포함합니다. 다음 예제에서는 HBase 사이트 구성 값을 `s3://amzn-s3-demo-bucket/my-hbase-site.xml` 파일에 지정된 값으로 설정합니다. 다음 명령을 입력하고 *myKey*를 EC2 키 페어의 이름으로 바꾸고 *amzn-s3-demo-bucket*을 Amazon S3 버킷의 이름으로 바꿉니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
        --applications Name=Hue Name=Hive Name=Pig Name=HBase \
        --use-default-roles --ec2-attributes KeyName=myKey \
        --instance-type c1.xlarge --instance-count 3 --termination-protected \
        --bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-
hbase,Args=["--site-config-file","s3://amzn-s3-demo-bucket/config.xml"]
```

HBase 작업을 사용자 지정하는 옵션을 두 개 이상 지정하는 경우 다음 예제에 표시된 대로 각 키-값 쌍 앞에 `-s` 옵션 스위치를 추가해야 합니다.

```
       --bootstrap-actions s3://elasticmapreduce/bootstrap-actions/configure-
hbase,Args=["-s","zookeeper.session.timeout=60000"]
```

프록시 세트와 SSH 연결이 열린 상태에서 http://*master-public-dns-name*:60010/master-status 가 포함된 브라우저 창을 열어서 HBase UI를 볼 수 있습니다. 여기서 *master-public-dns-name*은 HBase클러스터에 있는 마스터 노드의 퍼블릭 DNS 주소입니다.

SSH를 사용하여 마스터 노드에 연결하고 `mnt/var/log/hbase` 디렉터리로 이동하여 현재 HBase 로그를 볼 수 있습니다. 클러스터가 시작될 때 Amazon S3에 대한 로깅이 활성화되지 않은 경우 이 로그는 클러스터가 종료된 후 사용할 수 없습니다.

HBase 백업 및 복원

Amazon EMR은 수동 또는 자동 일정에 따라 HBase 데이터를 Amazon S3에 백업하는 기능을 제공합니다. 전체 및 증분 백업을 모두 수행할 수 있습니다. HBase 데이터의 백업 버전이 있으면 해당 버전을 HBase 클러스터로 복원할 수 있습니다. 현재 실행중인 HBase 클러스터로 복원하거나 백업된 데이터로 미리 채워진 새 클러스터를 시작할 수 있습니다.

백업 프로세스 중에 HBase는 쓰기 명령을 계속 실행합니다. 이렇게 하면 백업 중에 클러스터를 계속 사용할 수 있지만 백업 중인 데이터와 모든 쓰기 작업을 동시에 실행하는 데 불일치가 발생할 수 있습니다. 발생할 수 있는 불일치를 이해하려면 HBase가 클러스터의 노드에서 쓰기 작업을 분산시키는 것을 고려해야 합니다. 특정 노드가 폴링된 후 쓰기 작업이 발생하면 해당 데이터는 백업 아카이브에 포함되지 않습니다. 이미 폴링된 노드로 전송된 HBase 클러스터에 대한 이전 쓰기는 백업 아카이브에 없을 수 있으며 나중에 쓰기(폴링되기 전에 노드에 전송됨)가 포함됩니다.

이미 폴링된 노드로 전송된 HBase 클러스터에 대한 이전 쓰기는 백업 아카이브에 없을 수 있으며 나중에 쓰기(폴링되기 전에 노드에 전송됨)가 포함됩니다. 백업을 요청할 때 `--consistent` 파라미터를 지정하여 이 작업을 수행할 수 있습니다. 이 파라미터를 사용하면 이 기간 동안의 쓰기가 대기 상태가 되고 동기화가 완료되는 즉시 실행됩니다. 하나의 백업 단계에서 누락된 데이터가 다음 단계에서 백업되므로 시간 경과에 따른 불일치를 해결하는 반복 백업을 예약할 수도 있습니다.

HBase 데이터를 백업할 때 각 클러스터마다 다른 백업 디렉터리를 지정해야 합니다. 이를 수행하는 쉬운 방법은 클러스터 식별자를 백업 디렉터리에 지정된 경로의 일부로 사용하는 것입니다. 예를 들어 `s3://amzn-s3-demo-bucket/backups/j-3AEXXXXXX16F2`입니다. 이렇게 하면 향후 증분 백업이 올바른 HBase 클러스터를 참조하도록 보장할 수 있습니다.

더 이상 필요 없는 이전 백업 파일을 삭제할 준비가 되면 HBase 데이터의 전체 백업을 먼저 수행하는 것이 좋습니다. 이렇게 하면 모든 데이터가 보존되고 향후 증분 백업에 대한 기준이 제공됩니다. 전체 백업이 완료되면 백업 위치로 이동하여 수동으로 이전 백업 파일을 삭제할 수 있습니다.

HBase 백업 프로세스는 복사 작업에 S3DistCp를 사용합니다. 이 작업에는 임시 파일 스토리지 공간과 관련하여 특정 제한이 있습니다.

콘솔을 사용하여 HBase 백업 및 복원

콘솔은 새 클러스터를 시작하고 이전 HBase 백업의 데이터로 채울 수 있는 기능을 제공합니다. 또한 HBase 데이터를 주기적으로 증분 백업하도록 예약할 수 있습니다. 이미 실행 중인 클러스터로 데이터를 복원하고, 수동 백업을 수행하고, 자동화된 전체 백업을 예약하는 기능과 같은 추가 백업 및 복원 기능은 CLI를 통해 사용할 수 있습니다.

콘솔을 사용하여 아카이브된 HBase 데이터로 새 클러스터를 채우려면

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성을 선택합니다.
3. 소프트웨어 구성 섹션에서 추가 애플리케이션에 대해 HBase 및 구성 및 추가를 선택합니다.
4. 애플리케이션 추가 대화 상자에서 Restore From Backup(백업에서 복원)을 선택합니다.
5. 백업 위치에서 새 HBase 클러스터에 로드할 백업의 위치를 지정합니다. 이는 `s3://amzn-s3-demo-bucket/backups/` 양식의 Amazon S3 URL이어야 합니다.
6. Backup version(백업 버전)에 대해 값을 설정하여 로드할 백업 버전의 이름을 지정할 수 있습니다. 백업 버전의 값을 설정하지 않을 경우 Amazon EMR이 지정된 위치에 최신 백업을 로드합니다.
7. 추가를 선택하고 필요에 따라 다른 옵션으로 클러스터를 생성합니다.

콘솔을 사용하여 HBase 데이터의 자동 백업을 예약하려면

1. 소프트웨어 구성 섹션에서 추가 애플리케이션에 대해 HBase 및 구성 및 추가를 선택합니다.
2. Schedule Regular Backups(정기 백업 예약)를 선택합니다.
3. 백업의 일관성 여부를 지정합니다. 일관된 백업은 초기 백업 단계, 즉 노드 간 동기화 중에 쓰기 작업을 일시 중지하는 백업입니다. 이렇게 일시 중지된 쓰기 작업은 대기열에 배치되고 동기화가 완료되면 다시 시작됩니다.
4. Backup frequency(백업 빈도)에 숫자를 입력하고 일, 시간 또는 분을 선택하여 백업이 수행되는 빈도를 설정합니다. 실행되는 첫 번째 자동 백업은 전체 백업입니다. 그 후 Amazon EMR은 사용자가 지정한 일정에 따라 증분 백업을 저장합니다.
5. 백업이 저장되는 Amazon S3의 위치를 지정합니다. 증분 백업이 올바르게 계산되도록 하기 위해 각 HBase 클러스터를 Amazon S3의 별도 위치에 백업해야 합니다.
6. Backup Start Time(백업 시작 시간)에 대한 값을 설정하여 첫 번째 백업이 발생해야 할 시점을 지정합니다. 클러스터가 실행되는 즉시 첫 번째 백업이 시작되도록 now로 설정하거나, ISO 형식으

로 날짜와 시간을 입력할 수 있습니다. 예를 들어, 2013-09-26T20:00Z는 시작 시간을 2013년 9월 26일 오후 8시(UTC)로 설정합니다.

7. 추가를 선택합니다.
8. 계속해서 필요에 따라 다른 옵션으로 클러스터를 생성합니다.

CloudWatch를 이용한 HBase 모니터링

Amazon EMR은 HBase 백업을 모니터링하는 데 사용할 수 있는 세 가지 지표를 CloudWatch에 보고합니다. 이러한 지표는 5분 간격으로 CloudWatch로 푸시되고, 무료로 제공됩니다.

지표	설명
HBaseBackupFailed	마지막 백업 실패 여부입니다. 이 값은 기본적으로 0으로 설정되고 이전 백업 시도가 실패하면 1로 업데이트됩니다. 이 지표는 HBase 클러스터에 대해서만 보고됩니다. 사용 사례: HBase 백업 모니터링 Units: Count
HBaseMostRecentBackupDuration	이전 백업이 완료되는 데 걸린 시간입니다. 이 지표는 마지막으로 완료된 백업의 성공 여부에 관계없이 설정됩니다. 백업이 진행되는 동안 이 지표는 백업이 시작된 후의 시간(분)을 반환합니다. 이 지표는 HBase 클러스터에 대해서만 보고됩니다. 사용 사례: HBase 백업 모니터링 단위: 분
HBaseTimeSinceLastSuccessfulBackup	클러스터에서 마지막으로 성공한 HBase 백업이 시작된 후 경과된 시간 (분)입니다. 이 지표는 HBase 클러스터에 대해서만 보고됩니다. 사용 사례: HBase 백업 모니터링 단위: 분

HBase용 Ganglia 구성

configure-hbase-for-ganglia 부트스트랩 작업을 사용하여 HBase용 Ganglia를 구성합니다. 이 부트스트랩 작업은 HBase가 지표를 Ganglia에 게시하도록 구성합니다.

클러스터를 시작할 때 HBase와 Ganglia를 구성해야 합니다. 실행중인 클러스터에 Ganglia 보고를 추가할 수 없습니다.

또한 Ganglia는 `/mnt/var/log/ganglia/rrds`에서 서버에 로그 파일을 저장합니다. 로그 파일을 Amazon S3 버킷에 보관하도록 클러스터를 구성한 경우, Ganglia 로그 파일도 여기에 보관됩니다.

HBase용 Ganglia가 설치된 클러스터를 시작하려면 다음 예제에 표시된 대로 configure-hbase-for-ganglia 부트스트랩 작업을 사용합니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --ami-version 3.3 \
--applications Name=Hue Name=Hive Name=Pig Name=HBase Name=Ganglia \
--use-default-roles --ec2-attributes KeyName=myKey \
--instance-type c1.xlarge --instance-count 3 --termination-protected \
--bootstrap-actions Path=s3://elasticmapreduce/bootstrap-actions/configure-hbase-for-ganglia
```

Ganglia가 구성된 클러스터가 시작되면 마스터 노드에서 실행중인 그래픽 인터페이스를 사용하여 Ganglia 그래프와 보고서에 액세스할 수 있습니다.

Amazon EMR의 이전 AMI 버전에 고유한 Pig 애플리케이션

지원되는 Pig 버전

클러스터에 추가할 수 있는 Pig 버전은 사용하는 Amazon EMR AMI 버전 및 Hadoop 버전에 따라 다릅니다. 아래 표에서는 다양한 버전의 Pig와 호환되는 AMI 버전과 하둡 버전을 보여 줍니다. 사용 가능한 최신 버전의 Pig를 사용하여 성능 향상 및 새로운 기능을 활용하는 것이 좋습니다.

API를 사용하여 Pig를 설치할 때 RunJobFlow를 직접 호출하는 동안 클러스터에 Pig를 로드하는 단계의 인수로 `--pig-versions`를 지정하지 않으면 기본 버전이 사용됩니다.

Pig 버전	AMI 버전	구성 파라미터	Pig 버전 세부 정보

0.12.0
[릴리스 정보]
[설명서]
	3.1.0 이상	`--ami-version 3.1`
		`--ami-version 3.2`
		`--ami-version 3.3`
			다음에 대한 지원을 추가합니다.
			• JVM 구현 없는 UDF 스트리밍
			• ASSERT 및 IN 연산자
			• CASE 표현식
			• Pig 내장 함수인 AvroStorage.
			• 내장 함수인 ParquetLoader 및 ParquetStorer
			• BigInteger 및 BigDecimal 유형

0.11.1.1 [릴리스 정보] [설명서]	2.2 이상	`--pig-versions 0.11.1.1` `--ami-version 2.2`	입력이 Amazon S3에 상주하는 경우 PigStorage에서 LOAD 명령의 성능을 개선합니다.
0.11.1 [릴리스 정보] [설명서]	2.2 이상	`--pig-versions 0.11.1` `--ami-version 2.2`	JDK 7, Hadoop 2, Groovy 사용자 정의 함수, SchemaTuple 최적화, 새로운 연산자 등에 대한 지원을 추가합니다. 자세한 내용은 [Pig 0.11.1 change log]를 참조하세요.
0.9.2.2 [릴리스 정보]	2.2 이상	`--pig-versions 0.9.2.2`	하둡 1.0.3에 대한 지원을 추가합니다.

Pig 버전	AMI 버전	구성 파라미터	Pig 버전 세부 정보
설명서		`--ami-version 2.2`	
0.9.2.1 릴리스 정보 설명서	2.2 이상	`--pig-versions 0.9.2.1` `--ami-version 2.2`	MapR에 대한 지원을 추가합니다.
0.9.2 릴리스 정보 설명서	2.2 이상	`--pig-versions 0.9.2` `--ami-version 2.2`	여러 성능 향상과 버그 수정을 포함합니다. Pig 0.9.2 변경에 대한 전체 정보는 Pig 0.9.2 change log를 참조하세요.
0.9.1 릴리스 정보 설명서	2.0	`--pig-versions 0.9.1` `--ami-version 2.0`	
0.6 릴리스 정보	1.0	`--pig-versions 0.6` `--ami-version 1.0`	
0.3 릴리스 정보	1.0	`--pig-versions 0.3` `--ami-version 1.0`	

Pig 버전 세부 정보

Amazon EMR은 추가 Amazon EMR 패치가 적용되어 있을 수 있는 특정 Pig 릴리스를 지원합니다. Amazon EMR 클러스터에서 어떤 버전의 Pig를 실행할지를 구성할 수 있습니다. 이 작업을 수행하는 방법에 대한 자세한 내용은 Apache Pig 단원을 참조하세요. 다음 섹션에서는 다양한 Pig 버전과 Amazon EMR에서 로드된 버전에 적용되는 패치를 설명합니다.

Pig 패치

이 섹션에서는 Amazon EMR에서 사용할 수 있는 Pig 버전에 적용되는 사용자 지정 패치를 설명합니다.

Pig 0.11.1.1 패치

Amazon EMR의 Pig 0.11.1.1 버전은 입력이 Amazon S3에 상주하는 경우 PigStorage에서 LOAD 명령의 성능을 개선하는 유지 관리 릴리스입니다.

Pig 0.11.1 패치

Amazon EMR 버전의 Pig 0.11.1에는 Apache Software Foundation에서 제공하는 모든 업데이트와 Pig 버전 0.9.2.2 이후의 누적 Amazon EMR 패치가 포함됩니다. 하지만 Pig 0.11.1에는 새로운 Amazon EMR용 패치가 없습니다.

Pig 0.9.2 패치

Apache Pig 0.9.2는 Pig의 유지 관리 릴리스입니다. Amazon EMR 팀은 다음 패치를 Pig 0.9.2의 Amazon EMR 버전에 적용했습니다.

패치	설명
PIG-1429	부울 데이터 형식을 Pig에 첫 번째 클래스 데이터 형식으로 추가합니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-1429를 참조하십시오. 상태: 커밋됨 Apache Pig 버전에서 수정됨: 0.10
PIG-1824	Jython UDF에서 모듈 가져오기를 지원합니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-1824를 참조하십시오.

패치	설명
	상태: 커밋됨 Apache Pig 버전에서 수정됨: 0.10
PIG-2010	분산 캐시에 등록된 JAR을 번들링합니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-2010를 참조하십시오. 상태: 커밋됨 Apache Pig 버전에서 수정됨: 0.11
PIG-2456	사용자가 기본 Pig 문을 지정할 수 있는 ~/.pigbootup 파일을 추가합니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-2456를 참조하십시오. 상태: 커밋됨 Apache Pig 버전에서 수정됨: 0.11
PIG-2623	Amazon S3 경로를 사용하여 UDF를 등록하는 기능을 지원합니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-2623를 참조하십시오. 상태: 커밋됨 Apache Pig 버전에서 수정됨: 0.10, 0.11

Pig 0.9.1 패치

Amazon EMR 팀은 다음 패치를 Pig 0.9.1의 Amazon EMR 버전에 적용했습니다.

패치	설명
dfs에서 JAR 파일 및 Pig 스크립트 지원	HDFS, Amazon S3 또는 기타 분산 파일 시스템에 저장된 스크립트 실행 및 JAR 파일 등록에 대한 지원을 추가했습니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-1505를 참조하십시오.

패치	설명
	상태: 커밋됨 Apache Pig 버전에서 수정됨: 0.8.0
Pig에서 다중 파일 시스템 지원	한 파일 시스템에서 데이터를 읽고 다른 파일 시스템에 쓰는 Pig 스크립트에 대한 지원을 추가합니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-1564를 참조하십시오. 상태: 커밋 안 됨 Apache Pig 버전에서 수정됨: 해당 사항 없음
Piggybank 날짜/시간 및 문자열 UDF 추가	사용자 지정 Pig 스크립트를 지원하기 위해 날짜/시간 및 문자열 UDF를 추가합니다. 자세한 내용은 https://issues.apache.org/jira/browse/PIG-1565를 참조하십시오. 상태: 커밋 안 됨 Apache Pig 버전에서 수정됨: 해당 사항 없음

대화형 및 배치 Pig 클러스터

Amazon EMR을 사용하면 Pig 스크립트를 두 가지 모드로 실행할 수 있습니다.

- 대화형
- 배치

콘솔 또는를 사용하여 장기 실행 클러스터를 시작할 때를 사용하여 Hadoop 사용자로 마스터 노드ssh에 연결하고 Grunt 셸을 사용하여 대화형으로 Pig 스크립트를 개발하고 실행할 AWS CLI수 있습니다. Pig를 대화형으로 사용하면 배치 모드보다 더 쉽게 Pig 스크립트를 수정할 수 있습니다. 대화형 모드에서 Pig 스크립트를 수정한 후 스크립트를 Amazon S3에 업로드하고 배치 모드를 사용하여 프로덕션에서 스크립트를 실행할 수 있습니다. 또한 실행 중인 클러스터에서 대화형으로 Pig 명령을 제출하여 필요에 따라 분석하고 변환할 수 있습니다.

배치 모드에서 Pig 스크립트를 Amazon S3에 업로드한 다음, 작업을 클러스터에 하나의 단계로 제출합니다. 장기 실행 클러스터 또는 일시적 클러스터에 Pig 단계를 제출할 수 있습니다.

Amazon EMR의 이전 AMI 버전에 고유한 Spark 애플리케이션

대화식 또는 배치 모드로 Spark 사용

Amazon EMR에서는 두 가지 모드로 Spark 애플리케이션을 실행할 수 있습니다.

- 대화형
- 배치

콘솔 또는를 사용하여 장기 실행 클러스터를 시작할 때 SSH를 사용하여 Hadoop 사용자로 마스터 노드에 연결하고 Spark 셸을 사용하여 Spark 애플리케이션을 대화형으로 개발하고 실행할 AWS CLI수 있습니다. Spark를 대화식으로 사용하면 배치 환경에서보다 더 쉽게 Spark 애플리케이션을 시제품화하거나 테스트할 수 있습니다. Spark 애플리케이션을 대화형 모드에서 수정한 후에는 해당 애플리케이션 JAR 또는 Python 프로그램을 Amazon S3에 있는 클러스터의 로컬 프라이머리 노드 파일 시스템에 추가할 수 있습니다. 그런 다음 이 애플리케이션을 배치 워크플로로 제출할 수 있습니다.

배치 모드에서는 Spark 스크립트를 Amazon S3 또는 로컬 프라이머리 노드 파일 시스템에 업로드한 다음, 작업을 클러스터에 하나의 단계로 제출합니다. 장기 실행 클러스터 또는 일시적 클러스터에 Spark 단계를 제출할 수 있습니다.

Spark가 설치된 클러스터 생성

콘솔을 사용하여 Spark가 설치된 클러스터를 시작하려면

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성을 선택합니다.
3. 소프트웨어 구성에서 필요한 AMI 릴리스 버전을 선택합니다.
4. 설치할 애플리케이션의 목록에서 Spark를 선택한 다음 구성 및 추가를 선택합니다.
5. 인수를 추가하여 원하는 대로 Spark 구성을 변경합니다. 자세한 내용은 Spark 구성 단원을 참조하십시오. 추가(Add)를 선택합니다.
6. 필요에 따라 다른 옵션을 선택한 다음 클러스터 생성을 선택합니다.

다음은 Java를 사용하여 Spark가 설치된 클러스터를 생성하는 방법을 보여주는 예제입니다.

```
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
SupportedProductConfig sparkConfig = new SupportedProductConfig()
```

```
    .withName("Spark");

RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("Spark Cluster")
    .withAmiVersion("3.11.0")
    .withNewSupportedProducts(sparkConfig)
    .withInstances(new JobFlowInstancesConfig()
     .withEc2KeyName("myKeyName")
     .withInstanceCount(1)
     .withKeepJobFlowAliveWhenNoSteps(true)
     .withMasterInstanceType("m3.xlarge")
     .withSlaveInstanceType("m3.xlarge")
    );
RunJobFlowResult result = emr.runJobFlow(request);
```

Spark 구성

클러스터를 생성할 때 Github의 awslabs/emr-bootstrap-actions/spark 리포지토리에 있는 부트스트랩 작업을 실행하여 Spark를 구성합니다. 부트스트랩 작업이 허용하는 인수는 해당 리포지토리의 README를 참조하십시오. 부트스트랩 작업은 `$SPARK_CONF_DIR/spark-defaults.conf` 파일의 속성을 구성합니다. 설정에 대한 자세한 내용은 Spark 설명서의 Spark 구성 주제를 참조하십시오. 다음 URL의 "latest"를 `2.2.0` http://spark.apache.org/docs/latest/configuration.html과 같이 설치하는 Spark 버전 번호로 대체할 수 있습니다.

또한 각 애플리케이션 제출 시 Spark를 동적으로 구성할 수 있습니다. `spark` 구성 파일을 사용하여 실행기에 대한 리소스 할당을 자동으로 최대화하는 설정을 사용할 수 있습니다. 자세한 내용은 Spark 기본 구성 설정 재정의 단원을 참조하십시오.

Spark 기본 설정 변경

다음은 AWS CLI를 사용하여 `spark.executor.memory`가 2G로 설정된 클러스터를 생성하는 방법을 보여주는 예입니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Spark cluster" --ami-version 3.11.0 \
```

```
--applications Name=Spark, Args=[-d,spark.executor.memory=2G] --ec2-attributes
 KeyName=myKey \
--instance-type m3.xlarge --instance-count 3 --use-default-roles
```

Spark에 작업 제출

작업을 클러스터로 제출하려면 단계를 사용하여 EMR 클러스터에서 `spark-submit` 스크립트를 실행합니다. [AmazonElasticMapReduceClient](http://docs.aws.amazon.com/AWSJavaSDK/latest/javadoc/com/amazonaws/services/elasticmapreduce/AmazonElasticMapReduceClient.html)에서 `addJobFlowSteps` 메서드를 사용하여 단계를 추가합니다.

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
StepFactory stepFactory = new StepFactory();
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");

List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

StepConfig sparkStep = new StepConfig()
 .withName("Spark Step")
 .withActionOnFailure("CONTINUE")
 .withHadoopJarStep(stepFactory.newScriptRunnerStep("/home/hadoop/spark/bin/spark-
submit","--class","org.apache.spark.examples.SparkPi","/home/hadoop/spark/lib/spark-
examples-1.3.1-hadoop2.4.0.jar","10"));

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

Spark 기본 구성 설정 재정의

애플리케이션마다 Spark 기본 구성 값을 재정의해야 할 때가 있습니다. 이 작업은 단계를 사용하여 애플리케이션을 제출할 때 기본적으로 옵션을 `spark-submit`에 전달함으로써 수행할 수 있습니다. 예를 들면 `spark.executor.memory`를 변경하여 실행자 프로세스에 할당된 메모리를 변경해야 할 때가 있습니다. 다음과 같은 인수를 사용하여 `--executor-memory` 스위치를 공급할 수 있습니다.

```
/home/hadoop/spark/bin/spark-submit --executor-memory 1g --class
 org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/spark-examples*.jar 10
```

마찬가지로, `--executor-cores` 및 `--driver-memory`를 조정할 수 있습니다. 단계에서 다음 인수를 단계에 제공합니다.

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /home/hadoop/spark/lib/
spark-examples*.jar 10
```

또한 `--conf` 옵션을 사용하여 기본 제공 스위치가 없는 설정을 조정할 수도 있습니다. 조정 가능한 다양한 설정에 대한 자세한 내용은 Apache Spark 설명서의 [Dynamically Loading Spark Properties](https://spark.apache.org/docs/latest/configuration.html#dynamically-loading-spark-properties) 주제를 참조하세요.

Amazon EMR의 이전 AMI 버전에서 S3DistCp 유틸리티 차이

Amazon EMR에서 지원되는 S3DistCp 버전

Amazon EMR AMI 릴리스에서 다음 S3DistCp 버전이 지원됩니다. 1.0.7 이후 S3DistCp 버전은 클러스터에 있습니다. 최신 기능을 사용하려면 `/home/hadoop/lib`의 JAR을 사용하십시오.

버전	설명	릴리스 날짜
1.0.8	`--appendToLastFile` , `--requirePreviousManifest` 및 `--storageClass` 옵션을 추가합니다.	2014년 1월 3일
1.0.7	`--s3ServerSideEncryption` 옵션을 추가합니다.	2013년 5월 2일
1.0.6	`--s3Endpoint` 옵션을 추가합니다.	2012년 8월 6일
1.0.5	실행할 S3DistCp 버전을 지정하는 기능을 개선했습니다.	2012년 6월 27일
1.0.4	`--deleteOnSuccess` 옵션을 개선했습니다.	2012년 6월 19일
1.0.3	`--numberFiles` 및 `--startingIndex` 옵션에 대한 지원을 추가했습니다.	2012년 6월 12일
1.0.2	그룹 사용 시 파일 명명을 개선했습니다.	2012년 6월 6일
1.0.1	S3DistCp의 초기 릴리스.	2012년 1월 19일

클러스터에 S3DistCp 복사 단계 추가

S3DistCp 복사 단계를 실행 중인 클러스터에 추가하려면 다음 명령을 입력하고 *`j-3GYXXXXXX9IOK`*를 클러스터 ID로 바꾸고, *`amzn-s3-demo-bucket`*을 Amazon S3 버킷 이름으로 바꿉니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--s3Endpoint,s3-eu-west-1.amazonaws.com",\
"--src,s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node/",\
"--dest,hdfs:///output",\
"--srcPattern,.*[a-zA-Z,]+"]
```

Example Amazon CloudFront 로그를 HDFS에 로드

이 예제에서는 실행 중인 클러스터에 단계를 추가하여 Amazon CloudFront 로그를 HDFS에 로드합니다. 프로세스에서 Gzip(CloudFront 기본값)의 압축 형식을 LZO로 변경합니다. 이렇게 하면 LZO를 사용하여 압축된 데이터를 압축 해제 시 여러 맵으로 분할할 수 있으므로 Gzip 사용 시 압축이 완료될 때까지 기다릴 필요가 없어서 유용합니다. 따라서 Amazon EMR을 사용하여 데이터를 분석할 때 성능이 향상됩니다. 또한 이 예제는 `--groupBy` 옵션에 지정된 정규식을 사용하여 주어진 시간 동안 모든 로그를 단일 파일로 결합하여 성능을 개선합니다. Amazon EMR 클러스터는 여러 개의 작은 GZIP 압축 파일을 처리할 때보다 LZI로 압축된 소수의 대용량 파일을 처리할 때 더 효율적입니다. LZO 파일을 분할하려면 파일을 인덱싱하고 하둡-lzo 타사 라이브러리를 사용해야 합니다.

Amazon CloudFront 로그를 HDFS에 로드하려면 다음 명령을 입력하고 *j-3GYXXXXXX9IOK*를 클러스터 ID로 바꾸고, *amzn-s3-demo-bucket*을 Amazon S3 버킷 이름으로 바꿉니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK \
--steps Type=CUSTOM_JAR,Name="S3DistCp step",Jar=/home/hadoop/lib/emr-s3distcp-1.0.jar,
\
Args=["--src,s3://amzn-s3-demo-bucket/cf","--dest,hdfs:///local",\
"--groupBy,.*XABCD12345678.([0-9]+-[0-9]+-[0-9]+-[0-9]+).*",\
```

```
"--targetSize,128",
"--outputCodec,lzo","--deleteOnSuccess"]
```

이전 예제를 다음과 같은 CloudFront 로그 파일에서 실행한다고 가정합니다.

```
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.HLUS3JKx.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-01.I9CNAZrg.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.YRRwERSA.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.dshVLXFE.gz
s3://amzn-s3-demo-bucket/cf/XABCD12345678.2012-02-23-02.LpLfuShd.gz
```

S3DistCp는 파일을 복사 및 연결하고 다음 두 개 파일로 압축합니다. 이때 파일 이름은 정규식과 일치하는 항목에 따라 결정됩니다.

```
hdfs:///local/2012-02-23-01.lzo
hdfs:///local/2012-02-23-02.lzo
```

새로운 기능

이 페이지에서는 Amazon EMR 7.x, 6.x 및 5.x의 최신 릴리스에서 사용할 수 있는 변경 사항 및 기능을 설명합니다.

이러한 릴리스 정보는 각 릴리스의 애플리케이션 버전, 구성 요소 버전 및 사용 가능한 구성 분류와 함께 Amazon EMR 7.9.0, Amazon EMR 6.15.0 및 Amazon EMR 5.36.2 페이지에서도 사용할 수 있습니다.

- 이전 릴리스의 릴리스 정보는 Amazon EMR 릴리스 정보 아카이브 섹션을 참조하세요.
- 새로운 Amazon EMR 릴리스가 출시될 때 업데이트를 받으려면 Amazon EMR 릴리스 정보에 대한 RSS 피드를 구독하세요.

> **Note**
>
> Amazon EMR의 이후 릴리스에서는 AWS 서명 버전 4(SigV4)를 사용하여 Amazon S3에 대한 요청을 인증합니다. 새 S3 버킷에 액세스하고 워크로드가 중단되지 않도록 하려면 SigV4를 지원하는 Amazon EMR 릴리스를 사용하는 것이 좋습니다. 자세한 내용 및 SigV4를 지원하는 Amazon EMR 릴리스 목록은 Amazon EMR 및 AWS 서명 버전 4 섹션을 참조하세요.

Amazon EMR 7.9.0(7.x 시리즈의 최신 릴리스)

새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

다음 릴리스 정보에는 Amazon EMR 릴리스 7.9.0에 대한 정보가 포함되어 있습니다.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 7.9.0 application upgrades include AWS SDK for Java 2.31.16/1.12.782, Delta 3.3.0-amzn-1, Hadoop 3.4.1, Hudi 0.15.0-amzn-6, Iceberg 1.7.1-amzn-2, Spark 3.5.5, and Zookeeper 3.9.3.
- 영구 앱 UI에 대한 CM-CMK 지원 -이 기능은 EMR 서비스 버킷에서 관리되는 로그에 적용되는 SSE 암호화를 확장하여 고객 관리형 사용자 지정 마스터 키(CM-CMK)를 준수합니다. 이 기능을 사용하면 클러스터가 종료된 후 고객이 로그의 접근성을 관리할 수 있습니다.

알려진 문제

- EC2의 EMR에 대한 로깅 옵션이 비활성화된 경우 클러스터가 종료되고 영구 애플리케이션 UIs가 포함되지 않은 후에는 로그를 사용할 수 없습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Oozie 사용 중단 - OSS Oozie가 Attic 모드로 전환되었으므로 Apache Oozie를 포함하는 마지막 릴리스는 Amazon EMR 릴리스 7.9입니다. Apache Oozie는 Amazon EMR 릴리스 7.10 이상 릴리스에서 제거됩니다. Oozie 대신 MWAA를 사용하는 것이 좋습니다. 고객은 GitHub의 유틸리티 oozie-to-airflow-emr를 참조하여 기존 Oozie 워크플로를 MWAA로 마이그레이션할 수 있습니다.
- YARN 컨테이너 빈 패킹 - Amazon EMR 버전 7.9.0부터 YARN의 다중 노드 배치 정책을 기반으로 구축된 YARN 용량 스케줄러에 컨테이너 빈 패킹 정책을 사용할 수 있습니다.
- hudi-cli.sh - hudi-cli.sh 스크립트는 EMR 릴리스 7.9.0 이상에서 더 이상 사용되지 않습니다. Amazon EMR 릴리스 7.9.0 이상에서는 hudi-cli-bundle.jar을 사용합니다.
- FIPS 지원 AL2023 AMI의 EMR 클러스터 - EMR 7.9부터 고객은 FIPS 지원 AL2023 AMI에서 EMR 클러스터를 시작할 수 있습니다. 이 변경 사항은 모든 EMR 패키지가 FIPS 140-3에 대한 권장 암호화 정책을 준수하도록 EMR의 디지털 서명 및 해시 알고리즘을 업데이트합니다. 참고: EMR 또는 AL2023의 FIPS 인증 상태는 여기에 포함되지 않습니다. FIPS 140-3 검증 상태에 대한 최신 정보는 AL2023 FIPS FAQ를 참조하세요.
- 단계 지연 시간 -이 릴리스는 설치를 위해 하둡 애플리케이션을 명시적으로 선택하지 않고 전송 중 데이터 암호화를 사용할 때 단계 지연 시간이 증가하는 문제를 해결합니다.
- EBS 암호화 볼륨을 사용한 클러스터 시작 -이 릴리스에서는 EBS 암호화 볼륨에 대한 KMS 호출로 인해 EMR 클러스터 시작이 실패할 수 있는 문제를 해결합니다.

-

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 609.0일	6.1.140-154.222.amzn2023	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 20250521	6.1.134-152.225.amzn2023	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 7월 2025년 331.0	6.1.131-143.221.amzn2023	2025년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부), 아시아 태평양(도쿄)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 303.0	6.1.129-138.220.amzn2023	2025년 3월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 218.2일	6.1.128-136.201.amzn2023	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2023년 6월 2025년 211.0	6.1.127-135.201.amzn2023	2025년 2월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 아시아 태평양(말레이시아), 아시아 태평양(말레이시아), 멕시코(중부)

Amazon EMR 6.15.0(6.x 시리즈의 최신 릴리스)

새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

다음 릴리스 정보에는 Amazon EMR 릴리스 6.15.0에 대한 정보가 포함됩니다. 변경 사항은 6.14.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.15.0 변경 로그 섹션을 참조하세요.

새로운 특성

- 애플리케이션 업그레이드 – Amazon EMR 6.15.0 application upgrades include Apache Hadoop 3.3.6, Apache Hudi 0.14.0-amzn-0, Iceberg 1.4.0-amzn-0, and Trino 426.
- EC2에서 실행되는 EMR 클러스터의 실행 속도 향상 – 이제 EC2 클러스터에서 Amazon EMR을 실행하는 속도가 최대 35% 빨라졌습니다. 이번 개선 덕분에 고객은 대부분 대략 5분 만에 클러스터를 시작할 수 있게 되었습니다.
- EMR Studio용 CodeWhisperer – 이제 JupyterLab에서 코드를 작성할 때 Amazon CodeWhisperer와 Amazon EMR Studio를 함께 사용하면 실시간으로 권장 사항을 받을 수 있습니다. CodeWhisperer에서는 설명을 작성하고, 코드 한 줄을 완성하고, 라인별 권장 사항을 만들고, 완전한 형식의 함수를 생성할 수 있습니다.
- Flink를 사용한 더욱 빠른 작업 재시작 시간 – Amazon EMR 6.15.0 이상을 사용하면 태스크 복구 또는 규모 조정 작업 중에 작업 재시작 시간을 단축하기 위해 Apache Flink에 대해 이제 몇 가지 새로운 메커니즘을 사용할 수 있습니다. 이는 실행 그래프의 복구 및 재시작 속도를 최적화하여 작업 안정성을 향상시킵니다.
- 개방형 테이블 형식에 대한 테이블 수준 및 세분화된 액세스 제어 - Amazon EMR 6.15.0 이상에서는 Glue 데이터 카탈로그의 데이터에 액세스하는 EC2 클러스터의 Amazon EMR에서 Spark 작업을 실행할 때 AWS AWS Lake Formation 를 사용하여 Hudi, Iceberg 또는 Delta Lake 기반 테이블에 테이블, 행, 열 및 셀 수준 권한을 적용할 수 있습니다.
- Hadoop 업그레이드 - Amazon EMR 6.15.0에는 Apache Hadoop의 버전 3.3.6으로의 업그레이드가 포함되어 있습니다. Hadoop 3.3.6은 Amazon EMR 6.15가 배포되었을 시점에 최신 버전이었으며 Apache에서 2023년 6월에 릴리스했습니다. Amazon EMR의 이전 릴리스(6.9.0~6.14.x) 에서는 Hadoop 3.3.3이 사용되었습니다.

  업그레이드에는 재구성 가능한 데이터노드 파라미터, 모든 라이브 데이터노드에서 대량 재구성 작업을 시작하는 `DFSAdmin` 옵션, 탐색이 많은 리더가 읽을 여러 범위를 지정할 수 있도록 하는 벡터 API 등 수백 가지 개선 사항과 수정 사항이 포함되어 있습니다. Hadoop 3.3.6에서는 WAL(미리 쓰기 로그)에 대한 HDFS API 및 의미 체계에 대한 지원도 추가되어 HBase가 다른 스토리지 시스템 구현

에서 실행될 수 있습니다. 자세한 내용을 확인하려면Apache Hadoop 하둡 설명서에서 버전 3.3.4, 3.3.5 및 3.3.6의 변경 로그를 참조하세요.

- Java용 AWS SDK 버전 2 지원 - 애플리케이션이 v2를 지원하는 경우 Amazon EMR 6.15.0 애플리케이션은 Java용 AWS SDK 버전 1.12.569 또는 2.20.160을 사용할 수 있습니다. Java 2.x용 AWS SDK는 버전 1.x 코드 베이스의 주요 재작성입니다. Java 8+에 토대를 두고 있으며, 요청이 많았던 기능들을 몇 가지 추가했습니다. 여기에는 비차단 I/O에 대한 지원과 런타임에 다른 HTTP 구현을 연결하는 기능이 포함됩니다. Java용 SDK v1에서 v2로의 마이그레이션 가이드를 포함한 자세한 내용을 확인하려면 AWS Java용 SDK, 버전 2 가이드를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 TRUE (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 고가용성 EMR 클러스터를 개선하기 위해 이번 릴리스에서는 IPv6 엔드포인트를 사용하는 로컬 호스트의 Amazon EMR 대몬(daemon)에 대한 연결을 지원합니다.
- 이번 릴리스는 고가용성 클러스터의 모든 프라이머리 노드에 프로비저닝된 ZooKeeper와의 통신을 위한 TLS 1.2가 지원합니다.
- 이번 릴리스에서는 로그 파일이 범위를 초과하여 클러스터 작업을 방해하는 시나리오를 최소화하기 위해 프라이머리 노드에서 유지되는 ZooKeeper 트랜잭션 로그 파일의 관리를 개선합니다.
- 이번 릴리스는 고가용성 EMR 클러스터에 대한 노드 내 통신의 복원력을 강화합니다. 이러한 개선 사항은 부트스트랩 작업 실패나 클러스터 시작 실패 가능성을 줄입니다.
- Amazon EMR 6.15.0의 Tez는 Tez 그룹화된 분할에서 입력 분할을 비동기적으로 열도록 지정할 수 있는 구성을 도입합니다. 그 결과 단일 Tez 그룹화된 분할에 많은 수의 입력 분할이 있을 경우 읽기 쿼리 성능이 더욱 빨라집니다. 자세한 내용을 확인하려면 Tez 비동기식 분할 열기를 참조하세요.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2

릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Amzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 13일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

Amazon EMR 5.36.2(5.x 시리즈의 최신 릴리스)

새 Amazon EMR 릴리스는 최초 릴리스 날짜에 첫 번째 리전부터 시작하여 일정 기간에 걸쳐 여러 리전에 사용 가능해집니다. 이 기간에는 사용자 리전에서 최신 릴리스 버전이 제공되지 않을 수도 있습니다.

다음 릴리스 정보에는 Amazon EMR 릴리스 5.36.2에 대한 정보가 포함됩니다. 변경 사항은 5.36.1에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스에서는 Amazon EMR에서 클러스터의 HDFS 복제 인수 설정 미만으로 코어 노드를 스케일 다운하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이 개선 사항은 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 이 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항은 크기 조정 작업의 한 번 실패로 인해 클러스터 크기 조정이 무기한 실행될 위험을 줄여줍니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.
- Amazon EMR이 코어 노드를 정상적으로 해제하는 동안 클러스터 스케일 다운 작업이 중단되고 완전히 해제되기 전에 비정상 상태로 전환되는 문제를 수정합니다.
- Amazon EMR이 단일 노드를 재시작할 때 여러 프라이머리 노드가 있는 고가용성 클러스터의 노드 안정성을 개선합니다.
- Amazon EC2에서 실행되는 Amazon EMR을 사용하여 로그 관리를 최적화합니다. 따라서 클러스터 로그의 스토리지 비용이 조금 절감될 수 있습니다.
- 로그 파일이 범위를 초과하여 클러스터 작업을 방해하는 시나리오를 최소화하기 위해 프라이머리 노드에서 유지되는 ZooKeeper 트랜잭션 로그 파일의 관리를 개선합니다.
- Yarn ResourceManager와 통신할 수 없어 여러 프라이머리 노드가 있는 고가용성 클러스터에서 실패 원인이 될 수 있는 드문 버그를 수정합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 503.0	4.14.343	xxxxxx, 2024	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

Amazon EMR 및 AWS 서명 버전 4

Amazon EMR 릴리스는 AWS 서명 버전 4(SigV4)를 사용하여 Amazon S3에 대한 요청을 인증합니다. 2020년 6월 24일 이후 Amazon S3에서 생성된 버킷은 Signature Version 2(SigV2)로 서명된 요청을 지원하지 않습니다. 2020년 6월 24일 또는 이전에 생성된 버킷은 계속해서 SIGv2를 지원합니다. 새 S3 버킷에 액세스하고 워크로드가 중단되지 않도록 하려면 SigV4를 지원하는 Amazon EMR 릴리스로 마이그레이션하는 것이 좋습니다.

Amazon EMR에 포함된 애플리케이션(예: Apache Spark, Apache Hive, Presto)을 사용하는 경우 SigV4를 사용하도록 애플리케이션 코드를 변경하지 않아도 됩니다. Amazon EMR에 포함되지 않은 사용자 지정 애플리케이션을 사용하는 경우, SigV4를 사용하도록 코드를 업데이트해야 할 수 있습니다. 자세한 내용은 Amazon S3 사용 설명서에서 Signature Version 2에서 Signature Version 4로 전환을 참조하세요.

다음 Amazon EMR 릴리스는 SigV4를 지원합니다: emr-4.7.4, emr-4.8.5, emr-4.9.6, emr-4.10.1, emr-5.1.1, emr-5.2.3, emr-5.3.2, emr-5.4.1, emr-5.5.4, emr-5.6.1, emr-5.7.1, emr-5.8.3, emr-5.9.1, emr-5.10.1, emr-5.11.4, emr-5.12.3, emr-5.13.1, emr-5.14.2, emr-5.15.1, emr-5.16.1, emr-5.17.2, emr-5.18.1, emr-5.19.1, emr-5.20.1, emr-5.21.2, and emr-5.22.0 and higher 모든 6.x 및 7.x 릴리스에서 SigV4를 지원합니다.

CVE-2021-44228 완화를 위한 접근 방법

> **Note**
>
> Amazon EMR 릴리스 6.9.0 이상의 경우 Log4j 라이브러리를 사용하는 Amazon EMR에서 설치한 모든 구성 요소는 Log4j 버전 2.17.1 이상을 사용합니다.

EC2에서 실행되는 Amazon EMR

CVE-2021-44228에서 설명하는 문제는 신뢰할 수 없는 소스의 입력을 처리할 때 Apache Log4j 코어 버전(2.0.0에서 2.14.1 사이)과 관련이 있습니다. Amazon EMR 5.x 릴리스(최대 5.34.0) 및 EMR 6.x 릴리스(최대 Amazon EMR 6.5.0)에서 시작된 Amazon EMR 클러스터에는 이러한 버전의 Apache Log4j를 사용하는 Apache Hive, Flink, HUDI, Presto, Trino와 같은 오픈 소스 프레임워크가 포함되어 있습니다. 하지만 많은 고객이 Amazon EMR 클러스터에 설치된 오픈 소스 프레임워크를 사용하여 신뢰할 수 없는 소스의 입력을 처리하고 로깅합니다.

다음 섹션에 설명된 대로 'Log4j CVE-2021-44228에 대한 Amazon EMR 부트스트랩 작업 솔루션'을 적용하는 것이 좋습니다. 이 솔루션은 CVE-2021-45046 문제도 해결합니다.

> **Note**
>
> Amazon EMR의 부트스트랩 작업 스크립트는 Oozie의 증분 버그 수정 사항 및 개선 사항을 포함하도록 2022년 9월 7일에 업데이트되었습니다. Oozie를 사용하는 경우 다음 섹션에서 설명하는 업데이트된 Amazon EMR 부트스트랩 작업 솔루션을 적용해야 합니다.

Amazon EMR on EKS

기본 구성으로 Amazon EMR on EKS를 사용하는 경우 CVE-2021-44228에서 설명한 문제의 영향을 받지 않으므로 Log4j CVE-2021-44228 및 CVE-2021-45046에 대한 Amazon EMR 부트스트랩 작업 솔루션 섹션에 설명된 솔루션을 적용할 필요가 없습니다. Amazon EMR on EKS의 경우 Spark용 Amazon EMR 런타임은 Apache Log4j 버전 1.2.17을 사용합니다. Amazon EMR on EKS를 사용하는 경우 `log4j.appender` 구성 요소의 기본 설정을 `log`로 변경하면 안 됩니다.

Log4j CVE-2021-44228 및 CVE-2021-45046에 대한 Amazon EMR 부트스트랩 작업 솔루션

이 솔루션은 Amazon EMR 클러스터에 적용해야 하는 Amazon EMR 부트스트랩 작업을 제공합니다. 아래에서 각 Amazon EMR 릴리스에 대한 부트스트랩 작업 스크립트 링크를 제공합니다. 이 부트스트랩 작업을 적용하려면 다음 단계를 완료해야 합니다.

1. Amazon EMR 릴리스에 해당하는 스크립트를 AWS 계정의 로컬 S3 버킷에 복사합니다. Amazon EMR 릴리스에 맞는 부트스트랩 스크립트를 사용하는지 확인하세요.
2. EMR 설명서에서 설명한 지침에 따라 S3 버킷에 복사된 스크립트를 실행하도록 EMR 클러스터의 부트스트랩 작업을 설정합니다. EMR 클러스터에 대해 다른 부트스트랩 작업을 구성한 경우 이 스크립트가 실행할 첫 번째 부트스트랩 작업 스크립트로 설정되어 있는지 확인합니다.
3. 기존 EMR 클러스터를 종료하고 부트스트랩 작업 스크립트를 사용하여 새 클러스터를 시작합니다. AWS 테스트 환경에서 부트스트랩 스크립트를 테스트하고 프로덕션 환경에 적용하기 전에 애플리케이션을 검증하는 것이 좋습니다. EMR 마이너 릴리스의 최신 버전(예: 6.3.0)을 사용하지 않는 경우 최신 버전(예: 6.3.1)을 사용하고 그 다음에 위에서 설명한 솔루션을 적용해야 합니다.

CVE-2021-44228 및 CVE-2021-45046 - Amazon EMR 릴리스에 대한 부트스트랩 스크립트

Amazon EMR 릴리스 번호	스크립트 위치	스크립트 릴리스 날짜
6.5.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.5.0-v2.sh`	2022년 3월 24일
6.4.0	`s3://elasticmapreduce/bootstrap-actions/`	2022년 3월 24일

Amazon EMR 릴리스 번호	스크립트 위치	스크립트 릴리스 날짜
	`log4j/patch-log4j-emr-6.4.0-v2.sh`	
6.3.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.3.1-v2.sh`	2022년 3월 24일
6.2.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.2.1-v2.sh`	2022년 3월 24일
6.1.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.1.1-v2.sh`	2021년 12월 14일
6.0.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-6.0.1-v2.sh`	2021년 12월 14일
5.34.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.34.0-v2.sh`	2021년 12월 12일
5.33.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.33.1-v2.sh`	2021년 12월 12일

Amazon EMR 릴리스 번호	스크립트 위치	스크립트 릴리스 날짜
5.32.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.32.1-v2.sh`	2021년 12월 13일
5.31.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.31.1-v2.sh`	2021년 12월 13일
5.30.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.30.2-v2.sh`	2021년 12월 14일
5.29.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.29.0-v2.sh`	2021년 12월 14일
5.28.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.28.1-v2.sh`	2021년 12월 15일
5.27.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.27.1-v2.sh`	2021년 12월 15일
5.26.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.26.0-v2.sh`	2021년 12월 15일

Amazon EMR 릴리스 번호	스크립트 위치	스크립트 릴리스 날짜
5.25.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.25.0-v2.sh`	2021년 12월 15일
5.24.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.24.1-v2.sh`	2021년 12월 15일
5.23.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.23.1-v2.sh`	2021년 12월 15일
5.22.0	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.22.0-v2.sh`	2021년 12월 15일
5.21.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.21.2-v2.sh`	2021년 12월 15일
5.20.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.20.1-v2.sh`	2021년 12월 15일
5.19.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.19.1-v2.sh`	2021년 12월 15일

Amazon EMR 릴리스 번호	스크립트 위치	스크립트 릴리스 날짜
5.18.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.18.1-v2.sh`	2021년 12월 15일
5.17.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.17.2-v2.sh`	2021년 12월 15일
5.16.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.16.1-v2.sh`	2021년 12월 15일
5.15.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.15.1-v2.sh`	2021년 12월 15일
5.14.2	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.14.2-v2.sh`	2021년 12월 15일
5.13.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.13.1-v2.sh`	2021년 12월 15일
5.12.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.12.3-v2.sh`	2021년 12월 15일

Amazon EMR 릴리스 번호	스크립트 위치	스크립트 릴리스 날짜
5.11.4	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.11.4-v2.sh`	2021년 12월 15일
5.10.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.10.1-v2.sh`	2021년 12월 15일
5.9.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.9.1-v2.sh`	2021년 12월 15일
5.8.3	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.8.3-v2.sh`	2021년 12월 15일
5.7.1	`s3://elasticmapreduce/bootstrap-actions/log4j/patch-log4j-emr-5.7.1-v2.sh`	2021년 12월 15일

EMR 릴리스 버전	2021년 12월 기준 최신 개정
6.3.0	6.3.1
6.2.0	6.2.1
6.1.0	6.1.1
6.0.0	6.0.1

EMR 릴리스 버전	2021년 12월 기준 최신 개정
5.33.0	5.33.1
5.32.0	5.32.1
5.31.0	5.31.1
5.30.0 또는 5.30.1	5.30.2
5.28.0	5.28.1
5.27.0	5.27.1
5.24.0	5.24.1
5.23.0	5.23.1
5.21.0 또는 5.21.1	5.21.2
5.20.0	5.20.1
5.19.0	5.19.1
5.18.0	5.18.1
5.17.0 또는 5.17.1	5.17.2
5.16.0	5.16.1
5.15.0	5.15.1
5.14.0 또는 5.14.1	5.14.2
5.13.0	5.13.1
5.12.0, 5.12.1, 5.12.2	5.12.3
5.11.0, 5.11.1, 5.11.2, 5.11.3	5.11.4
5.9.0	5.9.1

EMR 릴리스 버전	2021년 12월 기준 최신 개정
5.8.0, 5.8.1, 5.8.2	5.8.3
5.7.0	5.7.1

자주 묻는 질문(FAQ)

- EMR 5 이전의 EMR 릴리스는 CVE-2021-44228의 영향을 받나요?

  아니요. EMR 릴리스 5 이전의 EMR 릴리스에서는 2.0 이전의 Log4j 버전을 사용합니다.
- 이 솔루션은 CVE-2021-45046 문제를 해결하나요?

  예. 이 솔루션은 CVE-2021-45046 문제도 해결합니다.
- 이 솔루션은 사용자가 EMR 클러스터에 설치한 사용자 지정 애플리케이션에도 적용되나요?

  부트스트랩 스크립트는 EMR에서 설치한 JAR 파일만 업데이트합니다. 사용자 지정 Amazon Linux AMI를 사용하거나 기타 메커니즘을 통해 EMR 클러스터에 제출한 단계로 부트스트랩 작업을 사용해 EMR 클러스터에서 사용자 지정 애플리케이션 및 JAR 파일을 설치하고 실행하는 경우 애플리케이션 공급업체에 문의하여 사용자 지정 애플리케이션이 CVE-2021- 44228의 영향을 받는지 확인하고 적절한 솔루션을 결정하세요.
- EMR on EKS에서 사용자 지정된 도커 이미지를 처리하려면 어떻게 해야 하나요?

  사용자 지정된 도커 이미지를 사용하여 Amazon EMR on EKS에 사용자 지정 애플리케이션을 추가하거나 사용자 지정 애플리케이션 파일을 사용하여 Amazon EMR on EKS에 작업을 제출하는 경우 애플리케이션 공급업체에 문의하여 사용자 지정 애플리케이션이 CVE-2021-44228의 영향을 받는지 확인하고 적절한 솔루션을 결정하세요.
- CVE-2021-44228 및 CVE-2021-45046에서 설명하는 문제를 완화하기 위해 부트스트랩 스크립트는 어떻게 작동하나요?

  부트스트랩 스크립트는 새 명령 세트를 추가하여 EMR 시작 명령을 업데이트합니다. 이 새 명령은 EMR에서 설치한 모든 오픈 소스 프레임워크에서 Log4j를 통해 사용되는 JndiLookup 클래스 파일을 삭제합니다. 이 명령은 Log4j 문제 해결을 위해 Apache에서 게시한 권장 사항을 따릅니다.
- Log4j 버전 2.17.1 이상을 사용하는 EMR 업데이트가 있나요?

  EMR 5 릴리스(최대 릴리스 5.34) 및 EMR 6 릴리스(최대 릴리스 6.5)에서는 최신 버전의 Log4j와 호환되지 않는 이전 버전의 오픈 소스 프레임워크를 사용합니다. 이러한 릴리스를 계속 사용하는 경우

부트스트랩 작업을 적용하여 CVE에서 논의된 문제를 완화하는 것이 좋습니다. EMR 5 릴리스 5.34 및 EMR 6 릴리스 6.5 이후에 Log4j 1.x 및 Log4j 2.x를 사용하는 애플리케이션은 각각 Log4j 1.2.17 이상 및 Log4j 2.17.1 이상을 사용하도록 업그레이드되었으며, CVE 문제를 완화하기 위해 부트스트랩 작업을 사용할 필요가 없습니다.

- EMR 릴리스는 CVE-2021-45105의 영향을 받나요?

  EMR의 기본 구성을 사용하여 Amazon EMR에서 설치한 애플리케이션은 CVE-2021-45105의 영향을 받지 않습니다. Amazon EMR에서 설치한 애플리케이션 중에서 Apache Hive만 컨텍스트 조회와 함께 Apache Log4j를 사용하며, 부적절한 입력 데이터를 처리할 수 있는 방식으로 기본이 아닌 패턴 레이아웃을 사용하지 않습니다.

- Amazon EMR은 다음 CVE 공개 사항의 영향을 받나요?

  다음 테이블에는 Log4j와 관련된 CVE 목록과 각 CVE가 Amazon EMR에 영향을 미치는지 여부가 나와 있습니다. 이 테이블의 정보는 Amazon EMR에서 기본 구성을 사용하여 애플리케이션을 설치하는 경우에만 적용됩니다.

CVE	EMR에 영향을 미치는지 여부	Notes
CVE-2022-23302	아니요	Amazon EMR은 Log4j JMSSink를 설치하지 않음
CVE-2022-23305	아니요	Amazon EMR은 Log4j JDbcpender를 설치하지 않음
CVE-2022-23307	아니요	Amazon EMR은 Log4j Chainsaw를 설치하지 않음
CVE-2020-9493	아니요	Amazon EMR은 Log4j Chainsaw를 설치하지 않음
CVE-2021-44832	아니요	Amazon EMR은 JNDI 연결 문자열을 사용하여 Log4j JDbcpender를 설치하지 않음
CVE-2021-4104	아니요	Amazon EMR은 Log4j JMPappender를 사용하지 않음

CVE	EMR에 영향을 미치는지 여부	Notes
CVE-2020-9488	아니요	Amazon EMR에서 설치한 애플리케이션은 Log4j SMTPPender를 사용하지 않음
CVE-2019-17571	아니요	Amazon EMR은 클러스터에 대한 퍼블릭 액세스를 차단하고 SocketServer를 시작하지 않음
CVE-2019-17531	아니요	최신 Amazon EMR 릴리스 버전으로 업그레이드하는 것이 좋습니다. Amazon EMR 5.33.0 이상에서는 jackson-databind 2.6.7.4 이상을 사용하고, EMR 6.1.0 이상에서는 jackson-databind 2.10.0 이상을 사용합니다. 이러한 버전의 jackson-databind는 CVE의 영향을 받지 않습니다.

Amazon EMR 릴리스 정보 아카이브

모든 Amazon EMR 릴리스 정보는 아래에 나와 있습니다. 각 릴리스의 포괄적인 릴리스 정보는 Amazon EMR 6.x 릴리스 버전, Amazon EMR 4.x 릴리스 버전, Amazon EMR 5.x 릴리스 버전 섹션을 참조하세요.

새로운 Amazon EMR 릴리스가 출시될 때 업데이트를 받으려면 Amazon EMR 릴리스 정보에 대한 RSS 피드를 구독하세요.

릴리스 6.14.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.14.0에 대한 정보가 포함됩니다. 변경 사항은 6.13.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.14.0 변경 로그 섹션을 참조하세요.

새로운 특성

- Amazon EMR 6.14.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-2, Flink 1.17.1, Iceberg 1.3.1, and Trino 422.
- Amazon EMR Managed Scaling은 Amazon EMR 6.14.0 이상에서 생성한 클러스터에 대해 `ap-southeast-3` 아시아 태평양(자카르타) 리전에서 정식 출시되었습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 폐기 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 6.14.0 릴리스는 Amazon EC2에서 실행되는 Amazon EMR을 사용하여 로그 관리를 최적화합니다. 따라서 클러스터 로그의 스토리지 비용이 조금 절감될 수 있습니다.
- 6.14.0 릴리스에서는 Amazon EBS 볼륨에서 크기에 따른 많은 변형을 지원하는 여러 코어 인스턴스를 처리하도록 조정 워크플로를 개선합니다. 이 개선 사항은 코어 노드에만 적용되며, 태스크 노드의 스케일 다운 작업은 영향을 받지 않습니다.
- 6.14.0 릴리스에서는 Amazon EMR이 오픈 소스 애플리케이션(예: Apache Hadoop YARN ResourceManager and HDFS NameNode)과 상호 작용하는 방식을 개선합니다. 이 개선을 통해 클러스터 조정으로 인한 운영 지연 위험이 줄어들고 오픈 소스 애플리케이션과의 연결 문제로 인해 발생하는 시작 장애가 완화됩니다.
- 6.14.0 릴리스에서는 클러스터 시작 시 애플리케이션 설치를 최적화합니다. 이를 통해 특정 조합의 Amazon EMR 애플리케이션에 대한 클러스터 시작 시간이 개선됩니다.
- 6.14.0 릴리스에서는 사용자 지정 도메인이 있는 VPC에서 실행 중인 클러스터에서 코어 또는 태스크 노드가 다시 시작될 때 클러스터 스케일 다운 작업이 중단될 수 있는 문제를 수정합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2

릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용](#)을 참조하세요.

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부),이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 17일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023906.0	4.14.322	2023년 9월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

릴리스 6.13.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.13.0에 대한 정보가 포함됩니다. 변경 사항은 6.12.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.13.0 변경 로그 섹션을 참조하세요.

새로운 특성

- Amazon EMR 6.13.0 supports Apache Spark 3.4.1, Apache Spark RAPIDS 23.06.0-amzn-1, CUDA Toolkit 11.8.0, and JupyterHub 1.5.0.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 TRUE (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- 6.13.0 릴리스에서서는 클러스터 종료 명령 실행 시 모든 로그가 Amazon S3에 정기적으로 업로드되도록 Amazon EMR 로그 관리 대몬(daemon)을 개선합니다. 이를 통해 클러스터 종료 속도가 빨라집니다.
- 6.13.0 릴리스에서서는 모든 로그 파일을 Amazon S3에 일관되고 시기적절하게 업로드하도록 Amazon EMR 로그 관리 기능을 개선합니다. 이는 특히 장기 실행 EMR 클러스터에 유용합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용]을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

릴리스 6.12.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.12.0에 대한 정보가 포함됩니다. 변경 사항은 6.11.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.12.0 변경 로그 섹션을 참조하세요.

새로운 특성

- Amazon EMR 6.12.0 supports Apache Spark 3.4.0, Apache Spark RAPIDS 23.06.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.1-amzn-0, Apache Iceberg 1.3.0-amzn-0, Trino 414, and PrestoDB 0.281.
- Amazon EMR 릴리스 6.12.0 이상에서는 Apache Livy, HiveServer2(HS2)를 통한 Apache Hive, Trino, Presto, Hue와의 LDAP 통합을 지원합니다. 6.12.0 이상을 사용하는 EMR 클러스터에 Apache Spark 및 Apache Hadoop을 설치하고 LDAP를 사용하도록 구성할 수도 있습니다. 자세한 내용은 Amazon EMR에서 인증을 위해 Active Directory 또는 LDAP 서버 사용을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태ExecutorState라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정합니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Amazon EMR 릴리스 6.12.0 이상에서는 Flink에 대한 Java 11 런타임 지원을 제공합니다. 자세한 내용은 Java 11과 함께 실행하도록 Flink 구성 단원을 참조하십시오.
- 6.12.0 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항을 통해 크기 조정 작업의 한 번 실패로 인해 클러스터 크기 조정이 무기한 중단될 위험이 줄어듭니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.
- 6.12.0 릴리스는 정상적으로 사용 중지되는 코어 노드가 완전히 해제되기 전에 어떤 이유로든 비정상 상태가 되면 클러스터 스케일 다운 작업이 중단될 수 있는 문제를 수정합니다.
- 6.12.0 릴리스에서는 클러스터가 클러스터의 HDFS 복제 계수 설정 미만으로 코어 노드의 스케일 다운을 시도하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이는 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 6.12.0 릴리스에서는 인스턴스의 상태 변경을 로깅하는 속도를 높여 Amazon EMR에 대한 상태 모니터링 서비스의 성능과 효율성을 개선합니다. 이 개선 사항을 통해 여러 사용자 지정 클라이언트 도구 또는 타사 애플리케이션을 실행하는 클러스터 노드에서 성능이 저하될 가능성이 줄어듭니다.

- 6.12.0 릴리스에서는 Amazon EMR에 대한 클러스터 내 로그 관리 대몬(daemon)의 성능을 개선합니다. 따라서 동시성이 높은 단계를 실행하는 EMR 클러스터를 사용하면 성능이 저하될 가능성이 줄어듭니다.
- Amazon EMR 릴리스 6.12.0에서는 로컬 인스턴스 스토리지의 열린 파일 핸들 및 관련 프로세스에서 현재 사용 중인 모든 로그를 식별하도록 로그 관리 대몬(daemon)이 업그레이드되었습니다. 이 업그레이드를 통해 Amazon EMR은 로그를 Amazon S3에 아카이브한 후 파일을 올바르게 삭제하고 스토리지 공간을 확보할 수 있습니다.
- 6.12.0 릴리스에는 로컬 클러스터 파일 시스템에서 사용되지 않는 빈 단계 디렉터리를 삭제하는 로그 관리 대몬(daemon) 개선 사항이 포함되어 있습니다. 빈 디렉터리가 너무 많으면 Amazon EMR 대몬(daemon)의 성능이 저하되고 디스크가 과도하게 사용될 수 있습니다.
- 6.12.0 릴리스에서는 YARN 타임라인 서버 로그의 로그 로테이션을 지원합니다. 이를 통해 특히 장기 실행 클러스터의 경우 디스크 과다 사용 시나리오를 최소화할 수 있습니다.
- Amazon EMR 6.10.0 이상에서는 기본 루트 볼륨 크기가 15GB로 늘어났습니다. 이전 릴리스에서 기본 루트 볼륨 크기는 10GB였습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

릴리스 6.11.1

다음 릴리스 정보에는 Amazon EMR 릴리스 6.11.1에 대한 정보가 포함됩니다. 변경 사항은 6.11.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 6.11.1 변경 로그 섹션을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실

행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터의 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정합니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.

- 잠금 경합으로 인해 노드가 서비스 해제를 시도하면서 동시에 노드가 추가되거나 제거되면 노드가 교착 상태에 빠질 수 있습니다. 따라서 Hadoop Resource Manager(YARN)가 응답하지 않게 되고 수신 및 현재 실행 중인 모든 컨테이너에 영향을 미칩니다.
- 이 릴리스에는 다시 시작한 후 고가용성 클러스터를 장애 상태에서 복구할 수 있도록 지원하는 변경 사항이 포함되어 있습니다.
- 이 릴리스에는 Hue 및 HBase에 대한 보안 수정 사항이 포함되어 있습니다.
- 이 릴리스에서는 Amazon EMR을 사용하여 Spark에서 워크로드를 실행하는 클러스터가 `contains`, `startsWith`, `endsWith`, `like`에서 잘못된 결과를 자동으로 수신할 수 있는 문제를 수정했습니다. 이 문제는 Amazon EMR Hive3 Metastore Server(HMS)에 메타데이터가 있는 파티셔닝된 필드에서 표현식을 사용할 때 발생합니다.
- 이 릴리스에서는 사용자 정의 함수(UDF)가 없을 때 발생하는 Glue 측 제한 관련 문제를 수정했습니다.
- 이 릴리스에서는 YARN 서비스 해제 시 로그 푸셔가 컨테이너 로그를 S3로 푸시하기 전에 노드 로그 집계 서비스에서 컨테이너 로그를 삭제하는 문제를 수정했습니다.
- 이 릴리스에서는 Hadoop에 노드 레이블이 활성화된 경우 FairShare 스케줄러 지표와 관련된 문제를 수정했습니다.
- 이 릴리스에서는 `spark-defaults.conf`에서 `spark.yarn.heterogeneousExecutors.enabled` 구성의 기본 `true` 값을 설정할 때 Spark 성능에 영향을 주던 문제를 수정했습니다.
- 이 릴리스에서는 Reduce 작업이 셔플 데이터를 읽지 못하는 문제를 수정했습니다. 이 문제로 인해 Hive 쿼리가 실패하고 메모리 손상 오류가 발생했습니다.
- 이 릴리스에서는 Presto 또는 Trino를 실행하는 EMR 클러스터의 클러스터 조정 워크플로에 새로운 재시도 메커니즘을 추가합니다. 이 개선 사항을 통해 크기 조정 작업의 한 번 실패로 인해 클러스터 크기 조정이 무기한 중단될 위험이 줄어듭니다. 또한 클러스터를 더 빠르게 스케일 업 및 스케일 다운하므로 클러스터 활용도가 향상됩니다.

- 이 릴리스에서는 클러스터가 클러스터의 HDFS 복제 계수 설정 미만으로 코어 노드의 스케일 다운을 시도하지 않도록 클러스터 스케일 다운 로직을 개선합니다. 이는 데이터 중복성 요구 사항에 부합하며 조정 작업이 중단될 가능성을 줄여줍니다.
- 로컬 인스턴스 스토리지의 열린 파일 핸들 및 관련 프로세스에서 현재 사용 중인 모든 로그를 식별하도록 로그 관리 대몬(daemon)이 업그레이드되었습니다. 이 업그레이드를 통해 Amazon EMR은 로그를 Amazon S3에 아카이브한 후 파일을 올바르게 삭제하고 스토리지 공간을 확보할 수 있습니다.
- 이 릴리스에는 로컬 클러스터 파일 시스템에서 사용되지 않는 빈 단계 디렉터리를 삭제하는 로그 관리 대몬(daemon) 개선 사항이 포함되어 있습니다. 빈 디렉터리가 너무 많으면 Amazon EMR 대몬(daemon)의 성능이 저하되고 디스크가 과도하게 사용될 수 있습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 [Amazon EMR에 대한 기본 Amazon Linux AMI 사용](#)을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(말레이시아), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(말레이시아), 유럽(런던), , 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 멕시코(중부), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 아시아 태평양(타이베이), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(말레이시아), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(말레이시아), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브), 멕시코(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히), 아시아 태평양(말레이시아)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 캐나다 서부(캘거리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

릴리스 6.11.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.11.0에 대한 정보가 포함됩니다. 변경 사항은 6.10.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

새로운 특성

- Amazon EMR 6.11.0에서는 Apache Spark 3.3.2-amzn-0, Apache Spark RAPIDS 23.02.0-amzn-0, CUDA 11.8.0, Apache Hudi 0.13.0-amzn-0, Apache Iceberg 1.2.0-amzn-0, Trino 410-amzn-0, PrestoDB 0.279-amzn-0을 지원합니다.

변경 사항, 향상된 기능 및 해결된 문제

- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태ExecutorState라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 `TRUE` (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.
- Amazon EMR 6.11.0에서 DynamoDB 커넥터가 버전 5.0.0으로 업그레이드되었습니다. 버전 5.0.0은 를 사용합니다 AWS SDK for Java 2.x. 이전 릴리스는 AWS SDK for Java 1.x를 사용했습니다. 이번 업그레이드로, Amazon EMR 6.11에서 DynamoDB 커넥터를 사용하기 전에 코드를 테스트해 보는 것이 좋습니다.
- Amazon EMR 6.11.0용 DynamoDB 커넥터가 DynamoDB 서비스를 직접 호출할 때는 사용자가 `dynamodb.endpoint` 속성에 제공한 리전 값을 사용합니다. `dynamodb.endpoint` 사용 시 `dynamodb.region`도 구성하고, 두 속성이 모두 동일한 AWS 리전을 대상으로 지정하는 것이 좋습니다. `dynamodb.endpoint`를 사용하고 `dynamodb.region`를 구성하지 않으면 Amazon EMR 6.11.0용 DynamoDB 커넥터는 잘못된 리전 예외를 반환하고 Amazon EC2 인스턴스 메타데이터 서비스(IMDS)에서 AWS 리전 정보를 조정하려고 시도합니다. 커넥터가 IMDS에서 리전을 검색할 수 없는 경우 기본값은 미국 동부(버지니아 북부)(`us-east-1`)입니다. 다음 오류는 `dynamodb.region` 속성을 올바르게 구성하지 않을 경우 발생할 수 있는 잘못된 리전 예외의 예입니다. 2.x로의 AWS SDK for Java 업그레이드의 영향을 받는 클래스에 대한 `error software.amazon.awssdk.services.dynamodb.model.DynamoDbException: Credential should be scoped to a valid region.` 자세한 내용은 Amazon EMR - DynamoDB 커넥터의 GitHub 리포지토리에서 [1.x AWS SDK for Java 에서 2.x(#175)로 업그레이드](#) 커밋을 참조하세요.
- 이 릴리스에서는 열 이름 바꾸기 작업 후 Delta Lake를 사용하여 Amazon S3에 델타 테이블 데이터를 저장할 때 열 데이터가 `NULL`로 표시되는 문제를 수정했습니다. Delta Lake의 이 실험 기능에 대한 자세한 내용은 Delta Lake 사용 설명서에서 [Column rename operation](#)을 참조하세요.
- 6.11.0 릴리스에서는 여러 프라이머리 노드를 포함하는 클러스터에서 프라이머리 노드 중 하나를 복제하여 엣지 노드를 생성할 때 발생할 수 있는 문제를 수정했습니다. 복제된 엣지 노드로 인해 스케일 다운 작업이 지연되거나 프라이머리 노드의 메모리 사용률이 높아질 수 있습니다. EMR 클러스터

와 통신하기 위해 엣지 노드를 생성하는 방법에 대한 자세한 내용은 GitHub의 `aws-samples` 리포지토리에서 [Edge Node Creator](https://github.com/aws-samples)를 참조하세요.

- 6.11.0 릴리스에서는 Amazon EMR이 재부팅 후 Amazon EBS 볼륨을 인스턴스에 다시 마운트하는 데 사용하는 자동화 프로세스를 개선합니다.
- 6.11.0 릴리스에서는 Amazon EMR이 Amazon CloudWatch에 게시하는 Hadoop 지표에 간헐적으로 간격이 발생하는 문제를 수정했습니다.
- 6.11.0 릴리스에서는 클러스터의 노드 제외 목록이 포함된 YARN 구성 파일에 대한 업데이트가 디스크 과다 사용으로 인해 중단되는 EMR 클러스터 관련 문제를 수정했습니다. 업데이트가 불완전하면 향후 클러스터 스케일 다운 작업에 방해가 됩니다. 이 릴리스에서는 클러스터를 정상 상태로 유지하고 조정 작업을 예상대로 수행하도록 보장합니다.
- Amazon EMR 6.10.0 이상에서는 기본 루트 볼륨 크기가 15GB로 늘어났습니다. 이전 릴리스에서 기본 루트 볼륨 크기는 10GB였습니다.
- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications` 값이 `false`로 설정됩니다.

  이 수정 사항은 YARN-9608 때문에 발생한 문제를 해결하지만, Managed Scaling이 활성화된 클러스터에서는 셔플 데이터 손실로 인해 Hive 작업이 실패할 수 있습니다. 이번 릴리스에서는 Hive 워크로드에 대해 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data`도 설정하여 이러한 위험을 완화했습니다. 이 구성은 Amazon EMR 릴리스 6.11.0 이상에서만 사용할 수 있습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

> **Note**
>
> 이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인

하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, ListReleaseLabels API 또는 list-release-labels CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

릴리스 6.10.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.10.0에 대한 정보가 포함됩니다. 변경 사항은 6.9.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

새로운 특성

- Amazon EMR 6.10.0에서는 Apache Spark 3.3.1, Apache Spark RAPIDS 22.12.0, CUDA 11.8.0, Apache Hudi 0.12.2-amzn-0, Apache Iceberg 1.1.0-amzn-0, Trino 403, PrestoDB 0.278.1을 지원합니다.
- Amazon EMR 6.10.0에는 Hudi 테이블의 데이터에 대한 읽기 액세스를 제공하는 네이티브 Trino-Hudi 커넥터가 포함되어 있습니다. `trino-cli --catalog hudi`를 사용하여 커넥터를 활성화하고, `trino-connector-hudi`를 사용하여 요구 사항에 맞게 커넥터를 구성할 수 있습니다. Amazon EMR과의 기본 통합을 통해 더 이상 Hudi 테이블을 쿼리하는 데 `trino-connector-hive`를 사용하지 않아도 됩니다. 새 커넥터에서 지원되는 구성 목록은 Trino 설명서에서 Hudi connector 페이지를 참조하세요.
- Amazon EMR 릴리스 6.10.0 이상에서는 Apache Flink와의 Apache Zeppelin 통합을 지원합니다. 자세한 정보는 Amazon EMR의 Zeppelin에서 Flink 작업을 참조하세요.

알려진 문제

- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 6.10.0에서 이 문제를 해결하려면 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications`의 값을 `false`로 설정하면 됩니다. Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 기본적으로 구성이 `false`로 설정되어 있습니다.
- Spark 3.3.1(EMR 버전 6.10 이상에서 지원됨)부터 서비스 해제 호스트의 모든 실행기는 DECOMMISSIONING 상태`ExecutorState`라는 새 로 설정됩니다. Yarn은 폐기 중인 실행기를 사용하여 작업을 할당할 수 없으므로 필요한 경우 실행 중인 작업에 대해 새 실행기를 요청합니다. 따라서 EMR Managed Scaling, EMR Auto Scaling 또는 EMR-EC2 클러스터에서 사용자 지정 조정 메커니즘을 사용하는 동안 Spark DRA를 비활성화하면 Yarn이 각 작업에 허용되는 최대 실행기를 요청할 수 있습니다. 이 문제를 방지하려면 위의 기능 조합을 사용할 때 `spark.dynamicAllocation.enabled` 속성을 TRUE (기본값)로 설정한 상태로 둡니다. 또한 Spark 작업의 `spark.dynamicAllocation.maxExecutors` 및 `spark.dynamicAllocation.minExecutors` 속성에 대한 값을 설정하여 작업 실행 중에 할당된 실행기 수를 제한하여 최소 및 최대 실행기 제약 조건을 설정할 수도 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 6.10.0은 `minimal-json.jar`에서 Apache Spark용 Amazon Redshift 통합의 종속 항목을 제거하고 필요한 Spark-Redshift 관련 jar(`spark-redshift.jar`, `spark-avro.jar`, `RedshiftJDBC.jar`)을 Spark의 실행기 클래스 경로에 자동으로 추가합니다.
- 6.10.0 릴리스에서는 MR 클러스터의 추가 로그 폴더를 모니터링하도록 클러스터 내 로그 관리 대몬(daemon)이 개선되었습니다. 이 개선 사항은 디스크 과다 사용 시나리오를 최소화합니다.
- 6.10.0 릴리스에서는 클러스터 내 로그 관리 대몬(daemon)이 중지되면 해당 대몬(daemon)을 자동으로 다시 시작합니다. 이 개선 사항은 디스크 과다 사용으로 인해 노드가 비정상으로 보이는 위험을 줄입니다.
- Amazon EMR 6.10.0은 EMRFS 사용자 매핑에 대한 리전 엔드포인트를 지원합니다.
- Amazon EMR 6.10.0 이상에서는 기본 루트 볼륨 크기가 15GB로 늘어났습니다. 이전 릴리스에서 기본 루트 볼륨 크기는 10GB였습니다.
- 6.10.0 릴리스에서는 다른 모든 Spark 실행기가 YARN 리소스 관리자의 서비스 해제 호스트에 있을 때 해당 Spark 작업이 중단되는 문제를 수정했습니다.
- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 HIVE-20703 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 HIVE-22269 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

> **Note**
>
> 이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, `ListReleaseLabels` API 또는 `list-release-labels` CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsRelease Label(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(취리히), 유럽(밀라노), 유럽(스페인), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)
2.0.20230404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE)

릴리스 6.9.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.9.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 릴리스 6.8.0에 관련됩니다. 릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

새로운 기능

- Amazon EMR 릴리스 6.9.0에서는 Apache Spark RAPIDS 22.08.0, Apache Hudi 0.12.1, Apache Iceberg 0.14.1, Trino 398, Tez 0.10.2를 지원합니다.
- Amazon EMR 릴리스 6.9.0에는 새로운 오픈 소스 애플리케이션, Delta Lake 2.1.0이 포함되어 있습니다.

- Apache Spark용 Amazon Redshift 통합은 Amazon EMR 릴리스 6.9.0 이상에 포함되어 있습니다. 이전의 오픈 소스 도구였던, 이 기본 통합은 Spark 커넥터로, Amazon Redshift와 Amazon Redshift Serverless에서 데이터를 읽고 쓰는 Apache Spark 애플리케이션을 빌드할 수 있습니다. 자세한 내용은 Amazon EMR에서 Apache Spark용 Amazon Redshift 통합 사용 단원을 참조하십시오.
- Amazon EMR 릴리스 6.9.0은 클러스터 스케일 다운 중에 Amazon S3에 로그를 아카이브하기 위한 지원을 추가합니다. 이전에는 클러스터 종료 중에만 Amazon S3에 로그 파일을 아카이브할 수 있었습니다. 새로운 기능을 통해 클러스터에서 생성된 로그 파일은 노드가 종료된 후에도 Amazon S3에서 계속 유지됩니다. 자세한 내용은 클러스터 로깅 및 디버깅 구성을 참조하세요.
- 이제 Trino에는 장기 실행 쿼리를 지원하기 위해 내결함성 실행 메커니즘이 포함됩니다. 내결함성 실행은 실패한 쿼리 또는 구성 요소 작업을 재시도하여 쿼리 실패를 완화합니다.
- Iceberg, Kinesis 또는 Kafka와 같은 Flink 테이블 소스의 메타데이터 또는 Apache Hive 테이블의 통합된 `BATCH` 및 `STREAM` 처리를 위해 Amazon EMR에서 Apache Flink를 사용할 수 있습니다. AWS Management Console AWS CLI또는 Amazon EMR API를 사용하여 AWS Glue 데이터 카탈로그를 Flink의 메타스토어로 지정할 수 있습니다. 자세한 내용은 Amazon EMR에서 Flink 구성 단원을 참조하십시오.
- 이제 Amazon SageMaker AI Studio를 사용하여 Amazon EMR on EC2 클러스터에서 Apache Spark, Apache Hive 및 Presto 쿼리에 대한 AWS Identity and Access Management (IAM) 런타임 역할 및 AWS Lake Formation기반 액세스 제어를 지정할 수 있습니다. Amazon SageMaker 자세한 내용은 Amazon EMR 단계에 대한 런타임 역할 구성을 참조하세요.

알려진 문제

- Amazon EMR 릴리스 6.9.0의 경우 Trino는 Apache Ranger가 활성화된 클러스터에서 작동하지 않습니다. Ranger와 함께 Trino를 사용해야 하는 경우 지원에 문의하세요.
- Apache Spark용 Amazon Redshift 통합을 사용하고 Parquet 형식의 time, timetz, timestamp 또는 timestamptz(마이크로초 정밀도)를 사용하는 경우 커넥터는 시간 값을 가장 가까운 밀리초 값으로 반올림합니다. 해결 방법으로 텍스트 언로드 형식 `unload_s3_format` 파라미터를 사용합니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.

  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon SageMaker AI Studio에서 Amazon EMR 클러스터에 연결하면 403 금지 응답 코드와 함께 간헐적으로 실패할 수 있습니다. 이 오류는 클러스터에서 IAM 역할을 설정하는 데 60초를 초과하는 경우 발생합니다. 해결 방법으로 Amazon EMR 패치를 설치하여 재시도를 활성화하고 제한 시간을 최소 300초로 늘릴 수 있습니다. 다음 단계를 사용하여 클러스터를 시작할 때 부트스트랩 작업을 적용합니다.

  1. 다음 Amazon S3 URI에서 부트스트랩 스크립트와 RPM 파일을 다운로드합니다.

  ```
  s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/replace-rpms.sh
  s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/gcsc/emr-secret-agent-1.18.0-SNAPSHOT20221121212949.noarch.rpm
  ```

  2. 이전 단계의 파일을 사용자가 소유한 Amazon S3 버킷에 업로드합니다. 버킷은 클러스터를 시작하려는 AWS 리전 동일한에 있어야 합니다.

  3. EMR 클러스터를 시작할 때 다음 부트스트랩 작업을 포함합니다. *bootstrap_URI* 및 *RPM_URI*를 Amazon S3의 해당되는 URI로 바꿉니다.

  ```
  --bootstrap-actions "Path=bootstrap_URI,Args=[RPM_URI]"
  ```

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

- Apache Flink는 네이티브 S3 FileSystem 및 Hadoop FileSystem 커넥터를 제공합니다. 이 커넥터를 사용하면 애플리케이션에서 FileSink를 생성하고 Amazon S3에 데이터를 쓸 수 있습니다. 이 FileSink는 다음 두 가지 예외 중 하나로 인해 실패합니다.

```
java.lang.UnsupportedOperationException: Recoverable writers on Hadoop are only
 supported for HDFS
```

```
Caused by: java.lang.NoSuchMethodError:
 org.apache.hadoop.io.retry.RetryPolicies.retryOtherThanRemoteAndSaslException(Lorg/apache/hadoop/io/retry/RetryPolicy;Ljava/util/Map;)Lorg/apache/hadoop/io/retry/RetryPolicy;
                                                at
 org.apache.hadoop.yarn.client.RMProxy.createRetryPolicy(RMProxy.java:302) ~[hadoop-yarn-common-3.3.3-amzn-0.jar:?]
```

해결 방법으로 Amazon EMR 패치를 설치할 수 있습니다. 그러면 Flink에서 위 문제를 해결할 수 있습니다. 클러스터를 시작할 때 부트스트랩 작업을 적용하려면 다음 단계를 완료합니다.

1. flink-rpm을 사용자의 Amazon S3 버킷에 다운로드합니다. RPM 경로는 `s3://DOC-EXAMPLE-BUCKET/rpms/flink/`입니다.
2. 다음 URI를 사용하여 Amazon S3에서 부트스트랩 스크립트와 RPM 파일을 다운로드합니다. 를 클러스터를 시작하려는 AWS 리전 `regionName`로 바꿉니다.

   ```
   s3://emr-data-access-control-regionName/customer-bootstrap-actions/gcsc/replace-rpms.sh
   ```

3. Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. Amazon EMR 6.8.0

및 6.9.0에서 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

Amazon EMR 6.10.0에서 이 문제를 해결하려면 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications`의 값을 `false`로 설정합니다. Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 기본적으로 구성이 `false`로 설정되어 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 릴리스 6.9.0 이상의 경우 Log4j 라이브러리를 사용하는 Amazon EMR에서 설치한 모든 구성 요소는 Log4j 버전 2.17.1 이상을 사용합니다.
- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 6.9.0에서는 Spark SQL을 사용하여 데이터를 읽을 때 Apache Hudi를 사용한 Lake Formation 기반 액세스 제어에 대한 제한된 지원을 추가했습니다. Spark SQL을 사용하는 SELECT 쿼리가 지원되며 열 수준의 액세스 제어로 제한됩니다. 자세한 내용은 Hudi 및 Lake Formation을 참조하세요.
- Amazon EMR 6.9.0을 사용하여 노드 레이블을 활성화한 상태로 Hadoop 클러스터를 생성하면 YARN 지표 API에서 기본 파티션이 아닌 모든 파티션에서 집계된 정보를 반환합니다. 자세한 내용은 YARN-11414를 참조하세요.
- Amazon EMR 릴리스 6.9.0에서 Trino를 버전 398(Java 17 사용)로 업데이트했습니다. Amazon EMR 6.8.0에서 이전에 지원되던 Trino 버전은 Java 11에서 실행되는 Trino 388이었습니다. 이 변경 사항에 대한 자세한 내용은 Trino 블로그에서 Trino updates to Java 17을 참조하세요.
- 이 릴리스에서는 Apache BigTop 및 Amazon EMR on EC2 클러스터 시작 시퀀스 간 타이밍 시퀀스 불일치 문제를 수정했습니다. 이 타이밍 시퀀스 불일치는 시스템에서 둘 이상의 작업을 적절한 순서로 수행하는 대신 동시에 수행하려고 할 때 발생합니다. 따라서 특정 클러스터 구성에서 인스턴스 시작 제한 시간을 초과하고 클러스터 시작 시간이 느려지는 문제가 발생했습니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

> **Note**
>
> 이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, `ListReleaseLabels` API 또는 `list-release-labels` CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 캐나다 서부(캘거리), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 캐나다 서부(캘거리), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브), 캐나다 서부(캘거리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.2023404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 1월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2022 103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

릴리스 6.8.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.8.0에 대한 정보가 포함됩니다. 변경 사항은 6.7.0에 관련됩니다.

새로운 기능

- Amazon EMR 단계 기능은 이제 Apache Livy 엔드포인트와 JDBC 및 ODBC 클라이언트를 지원합니다. 자세한 내용은 Amazon EMR 단계에 대한 런타임 역할 구성을 참조하세요.
- Amazon EMR 릴리스 6.8.0은 Apache HBase 릴리스 2.4.12와 함께 제공됩니다. 이 HBase 릴리스에서는 HBase 테이블을 아카이브 및 삭제할 수 있습니다. Amazon S3 아카이브 프로세스는 모든 테이

블 파일의 이름을 아카이브 디렉터리로 바꿉니다. 이 프로세스는 비용이 많이 들고 시간이 오래 걸릴 수 있습니다. 이제 아카이브 프로세스를 건너뛰고 대용량 테이블을 빠르게 삭제할 수 있습니다. 자세한 내용은 HBase 쉘 사용 단원을 참조하십시오.

알려진 문제

- Hadoop 3.3.3에서는 애플리케이션이 완료될 때까지 컨테이너가 실행되는 노드를 서비스 해제 상태로 유지하는 YARN의 변경 사항(YARN-9608)을 도입했습니다. 이 변경으로 셔플 데이터와 같은 로컬 데이터가 손실되지 않으므로 작업을 다시 실행할 필요가 없습니다. Amazon EMR 6.8.0 및 6.9.0에서 이 접근 방식을 사용하면 Managed Scaling의 활성화 여부에 상관없이 클러스터에서 리소스 사용률이 낮아질 수도 있습니다.

  Amazon EMR 6.10.0에서 이 문제를 해결하려면 `yarn-site.xml`에서 `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-applications`의 값을 `false`로 설정합니다. Amazon EMR 릴리스 6.11.0 이상과 6.8.1, 6.9.1, 6.10.1에서는 이 문제를 해결하기 위해 기본적으로 구성이 `false`로 설정되어 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 쉘을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 릴리스 6.8.0은 Apache Hudi 0.11.1과 함께 제공되지만, Amazon EMR 6.8.0 클러스터는 Hudi 0.12.0의 오픈 소스 `hudi-spark3.3-bundle_2.12`와도 호환됩니다.
- Amazon EMR 릴리스 6.8.0은 Apache Spark 3.3.0과 함께 제공됩니다. 이번 Spark 릴리스에서는 Apache Log4j 2 및 `log4j2.properties` 파일을 사용하여 Spark 프로세스에서 Log4j를 구성합니다. 클러스터에서 Spark를 사용하거나 사용자 지정 구성 파라미터를 사용하여 EMR 클러스터를 생성하고 Amazon EMR 릴리스 6.8.0으로 업그레이드하려는 경우 Apache Log4j 2의 새로운 `spark-log4j2` 구성 분류 및 키 형식으로 마이그레이션해야 합니다. 자세한 내용은 Apache Log4j 1.x에서 Log4j 2.x로 마이그레이션 단원을 참조하십시오.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

Note

이 릴리스에서는 패치 릴리스가 1개 더 추가되었으므로 더 이상 자동 AMI 업데이트를 받지 않습니다. 패치 릴리스는 소수점 두 번째 자리 뒤의 숫자(`6.8.1`)로 표시됩니다. 최신 패치 릴리스를 사용하는지 확인하려면 릴리스 가이드에서 사용 가능한 릴리스를 확인하거나, 콘솔에서 클러스터를 생성할 때 Amazon EMR 릴리스 드롭다운을 확인하거나, `ListReleaseLabels` API 또는 `list-release-labels` CLI 작업을 사용합니다. 새 릴리스에 대한 업데이트를 받으려면 새로운 기능에서 RSS 피드를 구독하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 아시아 태평양(멜버른), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaLabel(Arzon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 아시아 태평양(멜버른), 이스라엘(텔아비브), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히), 아시아 태평양(멜버른), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아시아 태평양(멜버른), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20230404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230119.1	4.14.301	2023년 2월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221004.0	4.14.294	2022년 11월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220912.1	4.14.291	2022년 9월 6일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

알려진 문제

- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Spark 3.2.0에서는 기본적으로 `spark.hadoopRDD.ignoreEmptySplits`를 `true`로 설정하기 때문입니다. 해결 방법으로 `spark.hadoopRDD.ignoreEmptySplits`를 명시적으로 `false`로 설정합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.

- 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
- 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
  ```

릴리스 일정에 대한 자세한 내용은 변경 로그를 참조하세요.

릴리스 6.7.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.7.0에 대한 정보가 포함됩니다. 변경 사항은 6.6.0에 관련됩니다.

최초 릴리스 날짜: 2022년 7월 15일

새로운 기능

- Amazon EMR은 이제 Apache Spark 3.2.1, Apache Hive 3.1.3, HUDI 0.11, PrestoDB 0.272, Trino 0.378을 지원합니다.
- Amazon EMR on EC2 클러스터에 대한 EMR 단계(Spark, Hive)를 통해 IAM 역할 및 Lake Formation 기반 액세스 제어를 지원합니다.
- Apache Ranger 지원 클러스터에서 Apache Spark 데이터 정의 명령문을 지원합니다. 이제 Apache Ranger 지원 클러스터에서 Apache Hive 메타데이터를 읽고 쓰는 Trino 애플리케이션에 대한 지원이 포함됩니다. 자세한 내용은 [Enable federated governance using Trino and Apache Ranger on Amazon EMR](https://aws.amazon.com/blogs/big-data/enable-federated-governance-using-trino-and-apache-ranger-on-amazon-emr/)을 참조하세요.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스페인), 유럽(스톡홀름), 유럽(취리히), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 중동(UAE), 이스라엘(텔아비브), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤), 아시아 태평양(하이데라바드), 중동(UAE), 유럽(스페인), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024 131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelease Label(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20230404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230119.1	4.14.301	2023년 2월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221004.0	4.14.294	2022년 11월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220912.1	4.14.291	2022년 10월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.20220719.0	4.14.287	2022년 8월 10일	`us-west-1` , `eu-west-3` , `eu-north-1` , `ap-south-1` , `me-south-1`

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220606.1	4.14.281	2022년 7월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

알려진 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 셸을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 잘못된 `Hbase.compat.version`을 사용하므로 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Spark 3.2.0에서는 기본적으로 `spark.hadoopRDD.ignoreEmptySplits`를 `true`로 설정하기 때문입니다. 해결 방법으로 `spark.hadoopRDD.ignoreEmptySplits`를 명시적으로 `false`로 설정합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자](#)를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
  ```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

  ```
  --bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
  ```

- Amazon EMR 6.7 이하에서 실행되는 클러스터에서는 `GetClusterSessionCredentials` API가 지원되지 않습니다.

릴리스 6.6.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.6.0에 대한 정보가 포함됩니다. 변경 사항은 6.5.0에 관련됩니다.

최초 릴리스 날짜: 2022년 5월 9일

설명서 업데이트 날짜: 2022년 6월 15일

새로운 기능

- Amazon EMR 6.6에서는 Apache Spark 3.2, Apache Spark RAPIDS 22.02, CUDA 11, Apache Hudi 0.10.1, Apache Iceberg 0.13, Trino 0.367, PrestoDB 0.267을 지원합니다.
- Amazon EMR 5.36 이상, 6.6 이상 또는 7.0 이상의 최신 패치 릴리스가 적용된 클러스터를 시작하는 경우 Amazon EMR은 기본 Amazon EMR AMI에 대한 최신 Amazon Linux 2023 or Amazon Linux 2 릴리스를 사용합니다. 자세한 내용은 Amazon EMR에 대한 기본 Amazon Linux AMI 사용을 참조하세요.

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 610.0	4.14.355-277.647.amzn2	2025년 7월 14일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 중동(UAE), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 아시아 태평양(하이데라바드), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
			럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 이스라엘(텔아비브), 캐나다(중부), 유럽(스페인), 중국(닝샤), 유럽(취리히)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250527.1	4.14.355-277.647.amzn2	2025년 6월 19일	유럽(스톡홀름), 중동(바레인), 중국(베이징), 아시아 태평양(뭄바이), 유럽(파리), 아시아 태평양(자카르타), 미국 동부(오하이오), 아프리카(케이프타운), 유럽(아일랜드), 유럽(프랑크푸르트), 남아메리카(상파울루), 아시아 태평양(홍콩), 미국 동부(버지니아 북부), 아시아 태평양(서울), 아시아 태평양(오사카), 유럽(런던), 유럽(밀라노), 아시아 태평양(도쿄), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 미국 서부(오레곤), 미국 서부(캘리포니아 북부), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 캐나다(중부), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2025 512.0	4.14.355-277.643.amzn2	2025년 6월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250428.0	4.14.355-276.639.amzn2	2025년 5월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(밀라노), 유럽(파리), 유럽(스톡홀름), 아시아 태평양(홍콩), 아시아 태평양(자카르타), 아시아 태평양(뭄바이), 아시아 태평양(오사카), 아시아 태평양(서울), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(도쿄), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 동부), AWS GovCloud(미국 서부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250414.0	4.14.355-276.618.amzn2	2025년 5월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250321.0	4.14.355	2025년 4월 9일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250305.0	4.14.355	2025년 3월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250220.0	4.14.355	2025년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250201.0	4.14.355	2025년 2월 28일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250123.4	4.14.355	2025년 1월 27일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20250116.0	4.14.355	2025년 1월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20241217.0	4.14.355	2025년 1월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2024001.0	4.14.352	2024년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240816.0	4.14.350	2024년 8월 21일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240809.0	4.14.349	2024년 8월 20일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240719.0	4.14.348	2024년 7월 25일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240709.1	4.14.348	2024년 7월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240223.0	4.14.336	2024년 3월 8일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240131.0	4.14.336	2024년 2월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240124.0	4.14.336	2024년 2월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20240109.0	4.14.334	2024년 1월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 218.0	4.14.330	2024년 1월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231206.0	4.14.330	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231116.0	4.14.328	2023년 12월 11일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 101.0	4.14.327	2023년 11월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231020.1	4.14.326	2023년 11월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20231012.1	4.14.326	2023년 10월 26일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230926.0	4.14.322	2023년 10월 19일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오리건), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부), AWS GovCloud(미국 서부), AWS GovCloud(미국 동부), 중국(베이징), 중국(닝샤)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230906.0	4.14.322	2023년 10월 4일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230822.0	4.14.322	2023년 8월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230808.0	4.14.320	2023년 8월 24일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230727.0	4.14.320	2023년 8월 14일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(프랑크푸르트), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 캐나다(중부)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230719.0	4.14.320	2023년 8월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 유럽(스톡홀름), 유럽(밀라노), 유럽(스페인), 유럽(프랑크푸르트), 유럽(취리히), 유럽(아일랜드), 유럽(런던), 유럽(파리), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(하이데라바드), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아시아 태평양(자카르타), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인), 중동(UAE), 캐나다(중부), 이스라엘(텔아비브)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230628.0	4.14.318	2023년 7월 12일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230612.0	4.14.314	2023년 6월 23일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230504.1	4.14.313	2023년 5월 16일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230418.0	4.14.311	2023년 5월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 404.1	4.14.311	2023년 4월 18일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.2023 404.0	4.14.311	2023년 4월 10일	미국 동부(버지니아 북부), 유럽(파리)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230320.0	4.14.309	2023년 3월 30일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230307.0	4.14.305	2023년 3월 15일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20230207.0	4.14.304	2023년 2월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다 (중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양 (서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2023 119.1	4.14.301	2023년 2월 3일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221210.1	4.14.301	2023년 12월 22일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221103.3	4.14.296	2022년 12월 5일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20221004.0	4.14.294	2022년 11월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsRelea Label(Ar zon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.2022 912.1	4.14.291	2022년 10월 7일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)
2.0.2022 805.0	4.14.287	2022년 8월 30일	`us-west-1`

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220719.0	4.14.287	2022년 8월 10일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220426.0	4.14.281	2022년 6월 10일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

OsReleaseLabel(Amazon Linux 버전)	Amazon Linux 커널 버전	사용 가능한 날짜	지원되는 리전
2.0.20220406.1	4.14.275	2022년 5월 2일	미국 동부(버지니아 북부), 미국 동부(오하이오), 미국 서부(캘리포니아 북부), 미국 서부(오레곤), 캐나다(중부), 유럽(스톡홀름), 유럽(아일랜드), 유럽(런던), 유럽(파리), 유럽(프랑크푸르트), 유럽(밀라노), 아시아 태평양(홍콩), 아시아 태평양(뭄바이), 아시아 태평양(자카르타), 아시아 태평양(도쿄), 아시아 태평양(서울), 아시아 태평양(오사카), 아시아 태평양(싱가포르), 아시아 태평양(시드니), 아프리카(케이프타운), 남아메리카(상파울루), 중동(바레인)

- Amazon EMR 릴리스 6.6 이상에서 Log4j 1.x 및 Log4j 2.x를 사용하는 애플리케이션은 각각 Log4j 1.2.17 이상 및 Log4j 2.17.1 이상을 사용하도록 업그레이드되었으며, CVE 문제를 완화하기 위해 부트스트랩 작업을 사용할 필요가 없습니다.
- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.
- Amazon EMR 5.32.0 및 6.5.0부터 Apache Spark의 동적 실행기 크기 조정이 기본적으로 활성화됩니다. 이 기능을 켜거나 끄기 위해 `spark.yarn.heterogeneousExecutors.enabled` 구성 파라미터를 사용할 수 있습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR에서는 EMR 기본 AMI 옵션을 사용하고 Apache Hadoop, Apache Spark, Apache Hive와 같은 일반적인 애플리케이션만 설치하는 클러스터에 대해 클러스터 시작 시간을 평균 최대 80초까지 줄일 수 있습니다.

알려진 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 셸을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 잘못된 `Hbase.compat.version`을 사용하므로 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- Amazon EMR 버전 6.6.0, 6.7.0 및 6.8.0에서 Spark와 함께 DynamoDB 커넥터를 사용하는 경우 입력 분할이 비어 있지 않은 데이터를 참조하더라도 테이블에서 모든 읽기는 빈 결과를 반환합니다. Spark 3.2.0에서는 기본적으로 `spark.hadoopRDD.ignoreEmptySplits`를 `true`로 설정하기 때문입니다. 해결 방법으로 `spark.hadoopRDD.ignoreEmptySplits`를 명시적으로 `false`로 설정합니다. Amazon EMR 릴리스 6.9.0에서는 이 문제가 수정되었습니다.
- Trino 장기 실행 클러스터에서 Amazon EMR 6.6.0을 사용하면 Trino jvm.config의 가비지 수집 로깅 파라미터를 통해 가비지 수집 로그에서 더 나은 인사이트를 얻을 수 있습니다. 이 변경으로 많은 가비지 수집 로그가 launcher.log(/var/log/trino/launcher.log) 파일에 추가됩니다. Amazon EMR 6.6.0에서 Trino 클러스터를 실행하는 경우 클러스터를 며칠 동안 실행한 후에 추가된 로그로 인해 노드의 디스크 공간이 부족해질 수 있습니다.

  이 문제의 해결 방법은 Amazon EMR 6.6.0에서 클러스터를 생성하거나 복제하는 동안 아래 스크립트를 부트스트랩 작업으로 실행하여 jvm.config에서 가비지 수집 로깅 파라미터를 비활성화하는 것입니다.

  ```
  #!/bin/bash
    set -ex
    PRESTO_PUPPET_DIR='/var/aws/emr/bigtop-deploy/puppet/modules/trino'
    sudo bash -c "sed -i '/-Xlog/d' ${PRESTO_PUPPET_DIR}/templates/jvm.config"
  ```
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.

- 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

- Amazon EMR 릴리스 5.36.0 및 6.6.0~6.9.0까지는 Log4j2 속성의 잘못된 파일 이름 패턴 구성으로 인해 `SecretAgent` 및 `RecordServer` 서비스 구성 요소에 로그 데이터가 손실될 수 있습니다. 잘못된 구성으로 인해 구성 요소에서 하루에 하나의 로그 파일만 생성합니다. 로테이션 전략이 수행되면 예상대로 새 로그 파일을 생성하는 대신 기존 파일을 덮어씁니다. 해결 방법으로 부트스트랩 작업을 사용하여 매시간 로그 파일을 생성하고 파일 이름에 자동 증가 정수를 추가하여 교체를 처리합니다.

  Amazon EMR 6.6.0~6.9.0 릴리스의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-6x/replace-puppet.sh,Args=[]"
```

  Amazon EMR 5.36.0의 경우 클러스터를 시작할 때 다음 부트스트랩 작업을 사용합니다.

```
--bootstrap-actions "Path=s3://emr-data-access-control-us-east-1/customer-bootstrap-actions/log-rotation-emr-5x/replace-puppet.sh,Args=[]"
```

릴리스 5.35.0

Amazon EMR 릴리스 5.35.0의 릴리스 정보입니다.

다음 릴리스 정보에는 Amazon EMR 릴리스 5.35.0에 대한 정보가 포함됩니다. 변경 사항은 5.34.0에 관련됩니다.

최초 릴리스 날짜: 2022년 3월 30일

새로운 기능

- Log4j 1.x 및 Log4j 2.x를 사용하는 Amazon EMR 릴리스 5.35 애플리케이션은 각각 Log4j 1.2.17 이상 및 Log4j 2.17.1 이상을 사용하도록 업그레이드되었으며, 이전 릴리스의 CVE 문제를 완화하기 위해 부트스트랩 작업을 사용할 필요가 없습니다. CVE-2021-44228 완화를 위한 접근 방법을(를) 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

Flink 변경 사항

유형 변경	설명
업그레이드	• Flink 버전을 1.14.2로 업데이트합니다. • log4j를 2.17.1로 업그레이드했습니다.

Hadoop 변경 사항

유형 변경	설명
EMR 5.34.0 이후 Hadoop 오픈 소스 백포트	• YARN-10438: ClientRMService#getContainerReport()에서 null인 containerId 처리 • YARN-7266: 타임라인 서버 이벤트 핸들러 스레드 잠김 • YARN-10438: RollingLevelDb 파일이 손상되거나 누락된 경우 ATS 1.5가 시작되지 않음 • HADOOP-13500: 구성 속성 객체의 반복 동기화 • YARN-10651: AbstractYarnScheduler.updateNodeResource()에서 CapacityScheduler가 NPE와 충돌함 • HDFS-12221: XmlEditsVisitor에서 xerces 교체 • HDFS-16410: OfflineEditsXmlLoader에서 안전하지 않은 Xml 구문 분석

유형 변경	설명
Hadoop 변경 사항 및 수정 사항	• KMS 및 HttpFS에서 사용되는 Tomcat이 8.5.75로 업그레이드되었습니다. • FileSystemOptimizedCommitterV2에서 커미터를 생성하는 동안 정의된 commitJob 출력 경로에 성공 마커가 기록되었습니다. CommitJob과 작업 수준 출력 경로가 다를 수 있으므로 매니페스트 파일에 정의된 경로를 사용하도록 경로가 수정되었습니다. Hive 작업의 경우 동적 파티션 또는 UNION ALL과 같은 작업을 수행할 때 성공 마커가 올바르게 기록됩니다.

Hive 변경 사항

유형 변경	설명
Hive는 이러한 JIRA 수정 사항을 포함하여 오픈 소스 릴리스 2.3.9로 업그레이드되었습니다.	• HIVE-17155: HiveConf.java 내 findConfFile()의 conf 경로에 몇 가지 문제가 있음 • HIVE-24797: Avro 스키마를 구문 분석할 때 기본값 검증 비활성화 • HIVE-21563: registerAllFunctionsOnce를 비활성화하여 Table#getEmptyTable 성능 개선 • HIVE-18147: java.net.BindException에서 테스트에 실패할 수 있음: 주소가 이미 사용 중임 • HIVE-24608: Hive 2.3.x용 HMS 클라이언트에서 get_table로 다시 전환 • HIVE-21200: 벡터화 - parquet에 대해 java.lang.UnsupportedOperationException이 발생한 날짜 열 • HIVE-19228: commons-httpclient 3.x 사용 제거

유형 변경	설명
EMR 5.34.0 이후 Hive 오픈 소스 백포트	• HIVE-19990: 조인 조건에서 간격 리터럴을 사용하는 쿼리 실패 • HIVE-25824: branch-2.3을 log4j 2.17.0으로 업그레이드 • TEZ-4062: 태스크가 완료되면 추론적 시도 예약을 중단해야 함 • TEZ-4108: 추론적 실행 레이스 조건 중에 NullPointerException 발생 • TEZ-3918: tez.task.log.level 설정이 적용되지 않음
Hive 업그레이드 및 수정 사항	• Log4j 버전을 2.17.1로 업그레이드 • ORC 버전을 1.4.3으로 업그레이드 • ShuffleScheduler의 페널티 스레드로 인한 교착 상태 수정됨
새로운 특성	• AM 로그에 Hive 쿼리를 인쇄하는 기능이 추가되었습니다. 이 옵션은 기본적으로 비활성화되어 있습니다. 플래그/구성: `tez.am.emr.print.hive.query.in.log` . 상태(기본값): FALSE.

Oozie 변경 사항

유형 변경	설명
EMR 5.34.0 이후 Oozie 오픈 소스 백포트	• OOZIE-3652: NoSuchFileException 발생 시 Oozie 시작 관리자에서 디렉터리 나열을 재시도해야 함

Pig 변경 사항

유형 변경	설명
업그레이드	• log4j를 1.2.17로 업그레이드했습니다.

알려진 문제

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자]를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 5.34.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.34.0에 대한 정보가 포함됩니다. 변경 사항은 5.33.1에 관련됩니다.

최초 릴리스 날짜: 2022년 1월 20일

릴리스 날짜: 2022년 3월 21일

새로운 기능

- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그

래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.

- [Hudi] Hudi 구성을 단순화하기 위한 개선 사항. 기본적으로 낙관적 동시성 제어를 비활성화했습니다.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- 이전에는 다중 마스터 클러스터에서 리소스 관리자를 수동으로 다시 시작하면 Zookeeper와 같은 Amazon EMR 클러스터 내 대몬(daemon)이 Zookeeper znode 파일에서 이전에 해제되거나 손실된 노드를 모두 다시 로드했습니다. 이로 인해 특정 상황에서 기본 한도가 초과되었습니다. 이제 Amazon EMR은 Zookeeper 파일에서 1시간 넘게 경과된 해제되거나 손실된 노드 레코드를 제거하고 내부 한도를 늘렸습니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- Zeppelin이 버전 0.10.0으로 업그레이드되었습니다.
- Livy 수정 - 0.7.1로 업그레이드됨
- Spark 성능 개선 - EMR 5.34.0에서 특정 Spark 구성 값이 재정의되면 이기종 실행기가 비활성화됩니다.

- WebHDFS 및 HttpFS 서버는 기본적으로 비활성화됩니다. Hadoop 구성(`dfs.webhdfs.enabled`)을 사용하여 WebHDFS를 다시 활성화할 수 있습니다. HttpFS 서버는 `sudo systemctl start hadoop-httpfs`를 사용하여 시작할 수 있습니다.

알려진 문제

- Livy 사용자 가장과 함께 사용되는 Amazon EMR Notebooks 기능은 기본적으로 HttpFS 가 비활성화되었기 때문에 작동하지 않습니다. 이 경우 EMR 노트북은 Livy 가장 기능이 활성화된 클러스터에 연결할 수 없습니다. 해결 방법은 클러스터에 EMR 노트북을 연결하기 전에 `sudo systemctl start hadoop-httpfs`를 사용하여 HttpFS 서버를 시작하는 것입니다.
- Apache Hadoop HttpFS 서버는 기본적으로 비활성화되어 있기 때문에 Amazon EMR 6.4.0에서는 Hue 쿼리가 작동하지 않습니다. Amazon EMR 6.4.0에서 Hue를 사용하려면 `sudo systemctl start hadoop-httpfs`를 사용하여 Amazon EMR 프라이머리 노드에서 HttpFS 서버를 수동으로 시작하거나 Amazon EMR 단계를 사용합니다.
- Livy 사용자 위장과 함께 사용되는 Amazon EMR Notebooks 기능은 기본적으로 HttpFS 가 비활성화되었기 때문에 작동하지 않습니다. 이 경우 EMR 노트북은 Livy 가장 기능이 활성화된 클러스터에 연결할 수 없습니다. 해결 방법은 클러스터에 EMR 노트북을 연결하기 전에 `sudo systemctl start hadoop-httpfs`를 사용하여 HttpFS 서버를 시작하는 것입니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 6.5.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.5.0에 대한 정보가 포함됩니다. 변경 사항은 6.4.0에 관련됩니다.

최초 릴리스 날짜: 2022년 1월 20일

릴리스 날짜: 2022년 3월 21일

새로운 기능

- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.
- Amazon EMR 5.32.0 및 6.5.0부터 Apache Spark의 동적 실행기 크기 조정이 기본적으로 활성화됩니다. 이 기능을 켜거나 끄기 위해 `spark.yarn.heterogeneousExecutors.enabled` 구성 파라미터를 사용할 수 있습니다.
- 대규모 분석 데이터 세트를 위한 Apache Iceberg 오픈 테이블 형식에 대한 지원.
- ranger-trino-plugin 2.0.1-amzn-1 지원
- toree 0.5.0에 대한 지원

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR 6.5 릴리스 버전은 이제 Apache Iceberg 0.12.0을 지원하며, Apache Spark용 EMR 런타임, Presto용 Amazon EMR 런타임, Apache Hive용 Amazon EMR 런타임을 통해 런타임을 개선했습니다.
- Apache Iceberg는 Amazon S3에서 대형 데이터 세트를 위한 오픈 테이블 형식이며, 대형 테이블, 원자성 커밋, 동시 쓰기, SQL 호환 테이블 진화 등에서 빠른 쿼리 성능을 제공합니다. EMR 6.5에서 Iceberg 테이블 형식으로 Apache Spark 3.1.2를 사용할 수 있습니다.
- Apache Hudi 0.9에서는 Spark SQL DDL 및 DML 지원을 추가했습니다. 이를 통해 SQL 문만 사용하여 Hudi 테이블을 생성하고 업서트할 수 있습니다. Apache Hudi 0.9에는 쿼리 측 및 작성기 측 성능 개선도 포함되어 있습니다.
- Apache Hive용 Amazon EMR 런타임은 준비 작업 중에 이름 바꾸기 작업을 제거하여 Amazon S3의 Apache Hive 성능을 향상시키고, 테이블 복구에 사용되는 메타스토어 검사(MSCK) 명령의 성능을 개선합니다.

알려진 문제

- Amazon EMR 릴리스 6.5.0, 6.6.0 또는 6.7.0에서 Apache Spark 쉘을 통해 Apache Phoenix 테이블을 읽으면 Amazon EMR에서 잘못된 `Hbase.compat.version`을 사용하므로 `NoSuchMethodError`가 발생합니다. Amazon EMR 릴리스 6.8.0에서는 이 문제가 수정되었습니다.
- 고가용성(HA)의 Hbase 번들 클러스터에서 기본 볼륨 크기 및 인스턴스 유형으로 프로비저닝하지 못합니다. 이 문제의 해결 방법은 루트 볼륨 크기를 늘리는 것입니다.
- Apache Oozie에서 Spark 작업을 사용하려면 Oozie `workflow.xml` 파일에 다음 구성을 추가해야 합니다. 그렇지 않으면 Oozie가 시작하는 Spark 실행기의 클래스 경로에서 Hadoop 및 EMRFS와 같은 여러 중요한 라이브러리가 누락됩니다.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 6.4.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.4.0에 대한 정보가 포함됩니다. 변경 사항은 6.3.0에 관련됩니다.

최초 릴리스 날짜: 2021년 9월 20일

릴리스 날짜: 2022년 3월 21일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.12.31
- CloudWatch Sink 버전 2.2.0
- DynamoDB 커넥터 버전 4.16.0
- EMRFS 버전 2.47.0
- Amazon EMR Goodies 버전 3.2.0
- Amazon EMR Kinesis 커넥터 버전 3.5.0
- Amazon EMR Record Server 버전 2.1.0
- Amazon EMR 스크립트 버전 2.5.0
- Flink 버전 1.13.1
- Ganglia 버전 3.7.2
- AWS Glue Hive 메타스토어 클라이언트 버전 3.3.0
- Hadoop 버전 3.2.1-amzn-4
- HBase 버전 2.4.4-amzn-0
- HBase-operator-tools 1.1.0
- HCatalog 버전 3.1.2-amzn-5
- Hive 버전 3.1.2-amzn-5
- Hudi 버전 0.8.0-amzn-0
- Hue 버전 4.9.0
- Java JDK 버전 Corretto-8.302.08.1(빌드 1.8.0_302-b08)
- JupyterHub 버전 1.4.1
- Livy 버전 0.7.1-incubating
- MXNet 버전 1.8.0
- Oozie 버전 5.2.1
- Phoenix 버전 5.1.2
- Pig 버전 0.17.0
- Presto 버전 0.254.1-amzn-0
- Trino 버전 359
- Apache Ranger KMS(멀티 마스터 투명 암호화) 버전 2.0.0
- ranger-plugins 2.0.1-amzn-0

- ranger-s3-plugin 1.2.0
- SageMaker Spark SDK 버전 1.4.1
- Scala 버전 2.12.10(OpenJDK 64비트 서버 VM, Java 1.8.0_282)
- Spark 버전 3.1.2-amzn-0
- spark-rapids 0.4.1
- Sqoop 버전 1.4.7
- TensorFlow 버전 2.4.1
- tez 버전 0.9.2
- Zeppelin 버전 0.9.0
- Zookeeper 버전 3.5.7
- 커넥터 및 드라이버: DynamoDB 커넥터 4.16.0

새로운 특성

- [Managed Scaling] Spark 셔플 데이터 Managed Scaling 최적화 - Amazon EMR 버전 5.34.0 이상 및 EMR 버전 6.4.0 이상에서 Managed Scaling은 이제 Spark 셔플 데이터(Spark가 특정 작업을 수행하기 위해 파티션에 재분배하는 데이터)를 인식합니다. 셔플 작업에 대한 자세한 내용은 Spark 프로그래밍 안내서 및 Amazon EMR 관리 안내서에서 Amazon EMR에서 EMR Managed Scaling 사용을 참조하세요.
- Apache Ranger 지원 Amazon EMR 클러스터에서는 Apache Spark SQL을 사용하여 `INSERT INTO`, `INSERT OVERWRITE`, `ALTER TABLE`을 통해 Apache Hive 메타스토어 테이블에 데이터를 삽입하거나 업데이트할 수 있습니다. Spark SQL에서 ALTER TABLE을 사용하는 경우 파티션 위치는 테이블 위치의 하위 디렉터리여야 합니다. Amazon EMR은 현재 파티션 위치가 테이블 위치와 다른 파티션으로의 데이터 삽입을 지원하지 않습니다.
- PrestoSQL은 Trino로 이름이 바뀌었습니다.
- Hive: LIMIT 절에 언급된 레코드 수를 가져오는 즉시 쿼리 실행을 중지함으로써 LIMIT 절을 사용하는 간단한 SELECT 쿼리의 실행이 가속화됩니다. 간단한 SELECT 쿼리는 GROUP BY 및 ORDER BY 절이 없는 쿼리 또는 reducer 스테이지가 없는 쿼리입니다. 예를 들어 `SELECT * from <TABLE> WHERE <Condition> LIMIT <Number>`입니다.

Hudi 동시성 제어

- Hudi는 이제 낙관적 동시성 제어(OCC)를 지원합니다. 이 기능을 UPSERT 및 INSERT와 같은 쓰기 작업과 함께 활용하면 여러 작성기에서 동일한 Hudi 테이블로 변경할 수 있습니다. 이는 파일 수준

OCC이므로 변경 사항이 충돌하지 않는 한 두 커밋(또는 작성기)이 동일한 테이블에 쓸 수 있습니다. 자세한 내용은 Hudi concurrency control을 참조하세요.

- Amazon EMR 클러스터에는 Zookeeper가 설치되어 있고, 이를 OCC의 잠금 공급자로 활용할 수 있습니다. 이 기능을 더 쉽게 사용할 수 있도록 Amazon EMR 클러스터에는 다음 속성이 미리 구성되어 있습니다.

```
hoodie.write.lock.provider=org.apache.hudi.client.transaction.lock.ZookeeperBasedLockProvider
hoodie.write.lock.zookeeper.url=<EMR Zookeeper URL>
hoodie.write.lock.zookeeper.port=<EMR Zookeeper Port>
hoodie.write.lock.zookeeper.base_path=/hudi
```

OCC를 활성화하려면 Hudi 작업 옵에서 다음 속성을 구성하거나 Amazon EMR 구성 API를 사용하여 클러스터 수준에서 다음 속성을 구성해야 합니다.

```
hoodie.write.concurrency.mode=optimistic_concurrency_control
hoodie.cleaner.policy.failed.writes=LAZY (Performs cleaning of failed writes lazily
 instead of inline with every write)
hoodie.write.lock.zookeeper.lock_key=<Key to uniquely identify the Hudi table> (Table
 Name is a good option)
```

Hudi 모니터링: Hudi 지표를 보고하기 위한 Amazon CloudWatch 통합

- Amazon EMR은 Amazon CloudWatch에 Hudi 지표를 게시하는 기능을 지원합니다. 다음과 같은 필수 구성을 설정하여 이 기능을 활성화합니다.

```
hoodie.metrics.on=true
hoodie.metrics.reporter.type=CLOUDWATCH
```

- 변경할 수 있는 선택적 Hudi 구성은 다음과 같습니다.

설정	설명	값
hoodie.metrics.cloudwatch.report.period.seconds	Amazon CloudWatch에 지표를 보고하는 빈도(초)	기본값은 60초이며, Amazon CloudWatch에서 제공하는 기본 1분 분해능에 적합합니다.
hoodie.metrics.cloudwatch.metric.prefix	각 지표 이름에 추가할 접두사	기본값은 비어 있습니다(접두사 없음).

설정	설명	값
hoodie.metrics.cloudwatch.namespace	지표가 게시되는 Amazon CloudWatch 네임스페이스	기본값은 Hudi입니다.
hoodie.metrics.cloudwatch.maxDatumsPerRequest	Amazon CloudWatch에 대한 1개의 요청에 포함할 수 있는 최대 데이텀 수	기본값은 20이며, Amazon CloudWatch 기본값과 동일합니다.

Amazon EMR Hudi 구성 지원 및 개선 사항

- 이제 고객은 EMR 구성 API 및 재구성 기능을 활용하여 클러스터 수준에서 Hudi 구성을 설정할 수 있습니다. Spark, Hive 등과 같은 다른 애플리케이션과 마찬가지로 /etc/hudi/conf/hudi-defaults.conf를 통해 새로운 파일 기반 구성 지원이 도입되었습니다. EMR은 사용자 환경을 개선하기 위해 몇 가지 기본값을 구성합니다.

  - `hoodie.datasource.hive_sync.jdbcurl` 은 클러스터 Hive 서버 URL로 구성되며, 더 이상 지정하지 않아도 됩니다. 이전에 Amazon EMR 마스터 IP를 지정해야 했던 Spark 클러스터 모드에서 작업을 실행할 때 특히 유용합니다.

  - HBase 특정 구성으로, Hudi에서 HBase 인덱스를 사용하는 데 유용합니다.

  - Zookeeper 잠금 제공자 특정 구성으로, 동시성 제어에서 설명한 대로, 이를 통해 낙관적 동시성 제어(OCC)를 더 쉽게 사용할 수 있습니다.

- 통과해야 하는 구성 수를 줄이고 가능한 경우 자동으로 추론할 수 있도록 추가 변경 사항이 도입되었습니다.

  - `partitionBy` 키워드를 사용하여 파티션 열을 지정할 수 있습니다.

  - Hive Sync를 활성화하면 더 이상 `HIVE_TABLE_OPT_KEY`, `HIVE_PARTITION_FIELDS_OPT_KEY`, `HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY`를 전달하지 않아도 됩니다. 이러한 값은 Hudi 테이블 이름 및 파티션 필드에서 추론할 수 있습니다.

  - `KEYGENERATOR_CLASS_OPT_KEY`는 반드시 전달하지 않아도 되며, `SimpleKeyGenerator` 및 `ComplexKeyGenerator`의 보다 간단한 사례에서 추론할 수 있습니다.

Hudi Caveats

- Hudi는 Merge on Read(MoR) 및 부트스트랩 테이블에 대해 Hive에서 벡터화된 실행을 지원하지 않습니다. 예를 들어 `hive.vectorized.execution.enabled`가 true로 설정된 경우 Hudi 실시간 테이블의 `count(*)`에서 오류가 발생합니다. 해결 방법으로 `hive.vectorized.execution.enabled`를 `false`로 설정하여 벡터화된 읽기를 비활성화할 수 있습니다.
- 다중 작성기 지원은 Hudi 부트스트랩 기능과 호환되지 않습니다.
- Flink Streamer 및 Flink SQL은 이 릴리스의 실험 기능입니다. 이러한 기능은 프로덕션 배포에 사용하지 않는 것이 좋습니다.

변경 사항, 향상된 기능 및 해결된 문제

이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.

- 이전에는 다중 마스터 클러스터에서 리소스 관리자를 수동으로 다시 시작하면 Zookeeper와 같은 Amazon EMR 클러스터 내 대몬(daemon)이 Zookeeper znode 파일에서 이전에 해제되거나 손실된 노드를 모두 다시 로드했습니다. 이로 인해 특정 상황에서 기본 한도가 초과되었습니다. 이제 Amazon EMR은 Zookeeper 파일에서 1시간 넘게 경과된 해제되거나 손실된 노드 레코드를 제거하고 내부 한도를 늘렸습니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이

머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.

- Apache YARN 타임라인 서버 버전 1 및 1.5 성능 문제를 해결하도록 클러스터 구성

  Apache YARN 타임라인 서버 버전 1 및 1.5는 매우 자주 사용되는 대규모 EMR 클러스터, 특히 Amazon EMR의 기본 설정인 `yarn.resourcemanager.system-metrics-publisher.enabled=true`에서 성능 문제를 일으킬 수 있습니다. 오픈 소스 YARN 타임라인 서버 v2는 YARN 타임라인 서버 확장성과 관련된 성능 문제를 해결했습니다.

  이 문제의 다른 해결 방법은 다음과 같습니다.

  - yarn-site.xml 파일에서 yarn.resourcemanager.system-metrics-publisher.enabled=false를 구성합니다.
  - 아래 설명에 따라 클러스터를 생성할 때 이 문제를 해결하는 수정 사항을 활성화합니다.

  다음 Amazon EMR 릴리스에는 이 YARN 타임라인 서버 성능 문제에 대한 수정 사항이 포함되어 있습니다.

  EMR 5.30.2, 5.31.1, 5.32.1, 5.33.1, 5.34.x, 6.0.1, 6.1.1, 6.2.1, 6.3.1, 6.4.x

  위에 지정된 Amazon EMR 릴리스에서 수정 사항을 활성화하려면 [`aws emr create-cluster` 명령 파라미터](`--configurations file://./configurations.json`)를 사용하여 전달되는 구성 JSON 파일에서 이러한 속성을 `true`로 설정합니다. 또는 재구성 콘솔 UI를 사용하여 수정 사항을 활성화할 수도 있습니다.

  configurations.json 파일 내용 예제:

```
[
{
"Classification": "yarn-site",
"Properties": {
"yarn.resourcemanager.system-metrics-publisher.timeline-server-v1.enable-batch":
 "true",
"yarn.resourcemanager.system-metrics-publisher.enabled": "true"
},
"Configurations": []
}
]
```

- WebHDFS 및 HttpFS 서버는 기본적으로 비활성화됩니다. Hadoop 구성(`dfs.webhdfs.enabled`)을 사용하여 WebHDFS를 다시 활성화할 수 있습니다. HttpFS 서버는 `sudo systemctl start hadoop-httpfs`를 사용하여 시작할 수 있습니다.
- 이제 Amazon Linux 리포지토리에서 HTTPS가 기본적으로 활성화됩니다. Amazon S3 VPCE 정책을 사용하여 특정 버킷에 대한 액세스를 제한하는 경우 새 Amazon Linux 버킷 ARN `arn:aws:s3:::amazonlinux-2-repos-$region/*`을 정책에 추가해야 합니다(`$region`을 엔드포인트가 있는 리전으로 바꿈). 자세한 내용은 AWS 토론 포럼에서이 주제를 참조하세요. [공지 사항: 이제 Amazon Linux 2는 패키지 리포지토리에 연결하는 동안 HTTPS를 사용하는 기능을 지원합니다].
- Hive: 마지막 작업에 HDFS의 스크래치 디렉터리를 사용할 수 있게 함으로써 쓰기 쿼리 성능이 개선되었습니다. 최종 작업의 임시 데이터는 Amazon S3 대신 HDFS에 기록되며, 데이터는 Amazon S3 디바이스 사이가 아닌 HDFS에서 최종 테이블 위치(Amazon S3)로 이동되므로 성능이 향상됩니다.
- Hive: Glue 메타스토어 파티션 정리를 통해 쿼리 컴파일 시간이 최대 2.5배 개선되었습니다.
- 기본적으로 내장된 UDF를 Hive에서 Hive Metastore Server로 전달하면 Glue는 제한된 표현식 연산자만 지원하므로 내장된 UDF의 하위 세트만 Glue 메타스토어로 전달됩니다. `hive.glue.partition.pruning.client=true`를 설정하면 모든 파티션 정리가 클라이언트 측에서 수행됩니다. `hive.glue.partition.pruning.server=true`를 설정하면 모든 파티션 정리가 서버 측에서 수행됩니다.

알려진 문제

- Apache Hadoop HttpFS 서버는 기본적으로 비활성화되어 있기 때문에 Amazon EMR 6.4.0에서는 Hue 쿼리가 작동하지 않습니다. Amazon EMR 6.4.0에서 Hue를 사용하려면 `sudo systemctl start hadoop-httpfs`를 사용하여 Amazon EMR 프라이머리 노드에서 HttpFS 서버를 수동으로 시작하거나 [Amazon EMR 단계를 사용]합니다.
- Livy 사용자 위장과 함께 사용되는 Amazon EMR Notebooks 기능은 기본적으로 HttpFS 가 비활성화되었기 때문에 작동하지 않습니다. 이 경우 EMR 노트북은 Livy 가장 기능이 활성화된 클러스터에 연결할 수 없습니다. 해결 방법은 클러스터에 EMR 노트북을 연결하기 전에 `sudo systemctl start hadoop-httpfs`를 사용하여 HttpFS 서버를 시작하는 것입니다.
- Amazon EMR 버전 6.4.0에서 Phoenix는 Phoenix 커넥터 구성 요소를 지원하지 않습니다.
- Apache Oozie에서 Spark 작업을 사용하려면 Oozie `workflow.xml` 파일에 다음 구성을 추가해야 합니다. 그렇지 않으면 Oozie가 시작하는 Spark 실행기의 클래스 경로에서 Hadoop 및 EMRFS와 같은 여러 중요한 라이브러리가 누락됩니다.

```
<spark-opts>--conf spark.yarn.populateHadoopClasspath=true</spark-opts>
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 5.32.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.32.0에 대한 정보가 포함됩니다. 변경 사항은 5.31.0에 관련됩니다.

최초 릴리스 날짜: 2021년 1월 8일

업그레이드

- Amazon Glue 커넥터를 버전 1.14.0으로 업그레이드함
- Amazon SageMaker Spark SDK를 버전 1.4.1로 업그레이드함
- 버전 1.11.890 AWS SDK for Java 으로 업그레이드됨
- EMR DynamoDB Connector 버전 4.16.0을 업그레이드함
- EMRFS를 버전 2.45.0으로 업그레이드함
- EMR Log Analytics Metric을 버전 1.18.0으로 업그레이드함
- EMR MetricsAndEventsApiGateway 클라이언트를 버전 1.5.0으로 업그레이드함
- EMR Record Server를 버전 1.8.0으로 업그레이드함
- EMR S3 Dist CP를 버전 2.17.0으로 업그레이드함
- EMR Secret Agent를 버전 1.7.0으로 업그레이드함
- Flink를 버전 1.11.2로 업그레이드함

- Hadoop을 버전 2.10.1-amzn-0으로 업그레이드함
- Hive를 버전 2.3.7-amzn-3으로 업그레이드함
- Hue를 버전 4.8.0으로 업그레이드함
- Mxnet을 버전 1.7.0으로 업그레이드함
- OpenCV를 버전 4.4.0으로 업그레이드함
- Presto를 버전 0.240.1-amzn-0으로 업그레이드함
- Spark를 버전 2.4.7-amzn-0으로 업그레이드함
- TensorFlow를 버전 2.3.1로 업그레이드함

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.

- 업그레이드된 구성 요소 버전.
- 구성 요소 버전 목록은 이 안내서에서 Amazon EMR 릴리스 정보를 참조하세요.

새로운 특성

- Amazon EMR 5.32.0 및 6.5.0부터 Apache Spark의 동적 실행기 크기 조정이 기본적으로 활성화됩니다. 이 기능을 켜거나 끄기 위해 `spark.yarn.heterogeneousExecutors.enabled` 구성 파라미터를 사용할 수 있습니다.
- 인스턴스 메타데이터 서비스(IMDS) V2 지원 상태: Amazon EMR 5.23.1, 5.27.1 및 5.32 이상의 구성 요소는 모든 IMDS 직접 호출에 IMDSv2를 사용합니다. 애플리케이션 코드에서 IMDS를 직접 호출하는 경우, IMDSv1 및 IMDSv2를 모두 사용하거나, 추가 보안을 위해 IMDSv2만 사용하도록 IMDS를 구성할 수 있습니다. 다른 5.x EMR 릴리스의 경우 ImDSv1을 비활성화하면 클러스터 시작 오류가 발생합니다.
- Amazon EMR 5.32.0부터 Apache Ranger와 기본적으로 통합되는 클러스터를 시작할 수 있습니다. Apache Ranger는 Hadoop 플랫폼 전반에서 포괄적인 데이터 보안을 지원, 모니터링 및 관리할 수 있는 오픈 소스 프레임워크입니다. 자세한 내용은 Apache Ranger를 참조하세요. 기본 통합을 통해 자체 Apache Ranger를 이용해 Amazon EMR에서 세분화된 데이터 액세스 제어를 적용할 수 있습니다. Amazon EMR 릴리스 안내서에서 Amazon EMR과 Apache Ranger 통합을 참조하세요.
- Amazon EMR 릴리스 5.32.0은 Amazon EMR on EKS를 지원합니다. EMR on EKS를 시작하는 방법에 대한 자세한 내용은 Amazon EMR on EKS란 무엇인가요?를 참조하세요.
- Amazon EMR 릴리스 5.32.0은 Amazon EMR Studio(평가판)를 지원합니다. EMR Studio를 시작하는 방법에 대한 자세한 내용은 Amazon EMR Studio(평가판)을 참조하세요.
- 범위 관리형 정책: AWS 모범 사례에 맞게 Amazon EMR은 더 이상 사용되지 않을 정책을 대체하기 위해 v2 EMR 범위 기본 관리형 정책을 도입했습니다. Amazon EMR 관리형 정책을 참조하세요.

알려진 문제

- Amazon EMR 6.3.0 및 6.2.0 프라이빗 서브넷 클러스터의 경우 Ganglia 웹 UI에 액세스할 수 없습니다. '액세스 거부됨(403)' 오류가 발생합니다. Spark, Hue, JupyterHub, Zeppelin, Livy, Tez와 같은 다른 웹 UI는 정상적으로 작동합니다. 퍼블릭 서브넷 클러스터에서 Ganglia 웹 UI 액세스도 정상적으로 작동합니다. 이 문제를 해결하려면 `sudo systemctl restart httpd`를 사용하여 프라이머리 노드에서 httpd 서비스를 다시 시작합니다. 이 문제는 Amazon EMR 6.4.0에서 수정되었습니다.
- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최

대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`

2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
```

```
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon Linux 또는 Amazon Linux 2 Amazon Machine Image(AMI)를 실행하는 EMR 클러스터는 기본 Amazon Linux 동작을 사용하며, 재부팅이 필요한 중요한 핵심 커널 업데이트를 자동으로 다운로드하여 설치하지는 않습니다. 이는 기본 Amazon Linux AMI를 실행하는 다른 Amazon EC2 인스턴스와 동일한 동작입니다. Amazon EMR 릴리스가 출시된 후 재부팅이 필요한 새로운 Amazon Linux 소프트웨어 업데이트(예: 커널, NVIDIA, CUDA 업데이트)를 사용할 수 있게 되면 기본 AMI를 실행하는 EMR 클러스터 인스턴스는 해당 업데이트를 자동으로 다운로드하여 설치하지 않습니다. 커널 업데이트를 받으려면 최신 Amazon Linux AMI를 사용하도록 Amazon EMR AMI를 사용자 지정할 수 있습니다.

- AWS Ranger 통합 옵션을 지정하는 보안 구성을 생성하는 콘솔 지원은 현재 GovCloud 리전에서 지원되지 않습니다. 보안 구성은 CLI를 사용하여 수행할 수 있습니다. Amazon EMR 관리 안내서에서 EMR 보안 구성 생성을 참조하세요.

- Amazon EMR 5.31.0 또는 5.32.0을 사용하는 클러스터에서 AtRestEncryption 또는 HDFS 암호화를 활성화하면 Hive 쿼리에서 다음과 같은 런타임 예외가 발생합니다.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.

  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.

- 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자]를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 6.2.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.2.0에 대한 정보가 포함됩니다. 변경 사항은 6.1.0에 관련됩니다.

최초 릴리스 날짜: 2020년 12월 9일

최종 업데이트 날짜: 2021년 10월 4일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.11.828
- emr-record-server 버전 1.7.0
- Flink 버전 1.11.2
- Ganglia 버전 3.7.2
- Hadoop 버전 3.2.1-amzn-1
- HBase 버전 2.2.6-amzn-0
- HBase-operator-tools 1.0.0
- HCatalog 버전 3.1.2-amzn-0
- Hive 버전 3.1.2-amzn-3
- Hudi 버전 0.6.0-amzn-1
- Hue 버전 4.8.0
- JupyterHub 버전 1.1.0
- Livy 버전 0.7.0
- MXNet 버전 1.7.0

- Oozie 버전 5.2.0
- Phoenix 버전 5.0.0
- Pig 버전 0.17.0
- Presto 버전 0.238.3-amzn-1
- PrestoSQL 버전 343
- Spark 버전 3.0.1-amzn-0
- spark-rapids 0.2.0
- TensorFlow 버전 2.3.1
- Zeppelin 버전 0.9.0-preview1
- Zookeeper 버전 3.4.14
- 커넥터 및 드라이버: DynamoDB 커넥터 4.16.0

새로운 특성

- HBase: 커밋 단계에서 이름 바꾸기를 제거하고 지속적 HFile 추적을 추가했습니다. Amazon EMR 릴리스 안내서에서 지속적 HFile 추적을 참조하세요.
- HBase: Create a config that forces to cache blocks on compaction을 백포트했습니다.
- PrestoDB: 동적 파티션 정리 기능이 개선되었습니다. 규칙 기반 조인 재정렬은 파티셔닝되지 않은 데이터에서 작동합니다.
- 범위 관리형 정책: AWS 모범 사례에 맞게 Amazon EMR은 더 이상 사용되지 않을 정책을 대체하기 위해 v2 EMR 범위 기본 관리형 정책을 도입했습니다. Amazon EMR 관리형 정책을 참조하세요.
- 인스턴스 메타데이터 서비스(IMDS) V2 지원 상태: Amazon EMR 6.2 이상의 Amazon EMR 구성 요소는 모든 IMDS 직접 호출에 IMDSv2를 사용합니다. 애플리케이션 코드에서 IMDS를 직접 호출하는 경우, IMDSv1 및 IMDSv2를 모두 사용하거나, 추가 보안을 위해 IMDSv2만 사용하도록 IMDS를 구성할 수 있습니다. 이전 Amazon EMR 6.x 릴리스에서 IMDSv1을 비활성화하면 클러스터를 시작하지 못합니다.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했

습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.

- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Spark: Spark 런타임의 성능이 개선되었습니다.

알려진 문제

- Amazon EMR 6.2에서는 EMR 6.2.0의 /etc/cron.d/libinstance-controller-java 파일에 잘못된 권한이 설정되어 있습니다. 파일의 권한이 644(-rw-r--r--)여야 하는 경우 645(-rw-r--r-x)입니다. 따라서 Amazon EMR 버전 6.2는 인스턴스 상태 로그를 로깅하지 않으며 /emr/instance-logs 디렉터리는 비어 있습니다. 이 문제는 Amazon EMR 6.3.0 이상에서 수정되었습니다.

  이 문제를 해결하려면 클러스터 시작 시 다음 스크립트를 부트스트랩 작업으로 실행합니다.

  ```
  #!/bin/bash
  sudo chmod 644 /etc/cron.d/libinstance-controller-java
  ```

- Amazon EMR 6.2.0 및 6.3.0 프라이빗 서브넷 클러스터의 경우 Ganglia 웹 UI에 액세스할 수 없습니다. '액세스 거부됨(403)' 오류가 발생합니다. Spark, Hue, JupyterHub, Zeppelin, Livy, Tez와 같은 다른 웹 UI는 정상적으로 작동합니다. 퍼블릭 서브넷 클러스터에서 Ganglia 웹 UI 액세스도 정상적으

로 작동합니다. 이 문제를 해결하려면 `sudo systemctl restart httpd`를 사용하여 프라이머리 노드에서 httpd 서비스를 다시 시작합니다. 이 문제는 Amazon EMR 6.4.0에서 수정되었습니다.

- Amazon EMR 6.2.0에는 httpd에서 계속 실패하여 Ganglia를 사용할 수 없게 되는 문제가 있습니다. 'cannot connect to the server' 오류가 발생합니다. 이 문제를 포함하는 이미 실행 중인 클러스터를 수정하려면 클러스터 프라이머리 노드에 SSH로 연결하고 `/etc/httpd/conf/httpd.conf`에 있는 `httpd.conf` 파일에 `Listen 80` 줄을 추가합니다. 이 문제는 Amazon EMR 6.3.0에서 수정되었습니다.
- 보안 구성을 사용하는 경우 EMR 6.2.0 클러스터에서 HTTPD에 오류가 발생합니다. 이로 인해 Ganglia 웹 애플리케이션 사용자 인터페이스를 사용할 수 없습니다. Ganglia 웹 애플리케이션 사용자 인터페이스에 액세스하려면 클러스터의 프라이머리 노드에 있는 `/etc/httpd/conf/httpd.conf` 파일에 `Listen 80`을 추가합니다. 클러스터 연결에 대한 자세한 내용은 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.

  또한 보안 구성을 사용할 때 EMR Notebooks는 EMR 6.2.0 클러스터와의 연결을 설정하지 못합니다. 노트북은 커널을 나열하지 못하고 Spark 작업을 제출하지 못합니다. 대신 다른 Amazon EMR 버전에서 EMR Notebooks를 사용하는 것이 좋습니다.
- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

     ```
     LimitNOFILE=65536
     ```

```
LimitNPROC=65536
```

2. InstanceController 다시 시작

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- **Important**

  Amazon EMR 6.1.0 및 6.2.0에는 모든 Hudi 삽입, 업서트 및 삭제 작업에 큰 영향을 미칠 수 있는 성능 문제가 포함되어 있습니다. Amazon EMR 6.1.0 또는 6.2.0에서 Hudi를 사용하려는 경우 AWS 지원팀에 문의하여 패치된 Hudi RPM을 받아야 합니다.

- > **Important**
  >
  > Amazon Linux 또는 Amazon Linux 2 Amazon Machine Image(AMI)를 실행하는 EMR 클러스터는 기본 Amazon Linux 동작을 사용하며, 재부팅이 필요한 중요한 핵심 커널 업데이트를 자동으로 다운로드하여 설치하지는 않습니다. 이는 기본 Amazon Linux AMI를 실행하는 다른 Amazon EC2 인스턴스와 동일한 동작입니다. Amazon EMR 릴리스가 출시된 후 재부팅이 필요한 새로운 Amazon Linux 소프트웨어 업데이트(예: 커널, NVIDIA, CUDA 업데이트)를 사용할 수 있게 되면 기본 AMI를 실행하는 EMR 클러스터 인스턴스는 해당 업데이트를 자동으로 다운로드하여 설치하지 않습니다. 커널 업데이트를 받으려면 최신 Amazon Linux AMI를 사용하도록 Amazon EMR AMI를 사용자 지정할 수 있습니다.

- Amazon EMR 6.2.0 Maven 아티팩트는 게시되지 않습니다. 향후 Amazon EMR 릴리스와 함께 게시됩니다.
- HBase 스토어파일 시스템 테이블을 사용한 지속적 HFile 추적은 HBase 리전 복제 기능을 지원하지 않습니다. HBase 리전 복제에 대한 자세한 내용은 타임라인이 일치하는 고가용성 읽기를 참조하세요.
- Amazon EMR 6.x 및 EMR 5.x Hive 버킷팅 버전 차이

  EMR 5.x는 OOS Apache Hive 2를 사용하지만, EMR 6.x는 OOS Apache Hive 3을 사용합니다. 오픈 소스 Hive2는 버킷팅 버전 1을 사용하는 반면, 오픈 소스 Hive3은 버킷팅 버전 2를 사용합니다. Hive 2(EMR 5.x) 및 Hive 3(EMR 6.x)의 버킷팅 버전 차이는 Hive 버킷팅 해싱 기능이 다르다는 것을 의미합니다. 아래 예제를 참조하십시오.

  다음 표는 각각 EMR 6.x 및 EMR 5.x에서 생성된 예제입니다.

  ```
  -- Using following LOCATION in EMR 6.x
  CREATE TABLE test_bucketing (id INT, desc STRING)
  PARTITIONED BY (day STRING)
  CLUSTERED BY(id) INTO 128 BUCKETS
  LOCATION 's3://your-own-s3-bucket/emr-6-bucketing/';

  -- Using following LOCATION in EMR 5.x
  LOCATION 's3://your-own-s3-bucket/emr-5-bucketing/';
  ```

  EMR 6.x 및 EMR 5.x 모두에 동일한 데이터를 삽입합니다.

  ```
  INSERT INTO test_bucketing PARTITION (day='01') VALUES(66, 'some_data');
  INSERT INTO test_bucketing PARTITION (day='01') VALUES(200, 'some_data');
  ```

S3 위치를 확인하는 경우 EMR 6.x(Hive 3) 및 EMR 5.x(Hive 2) 간에 해싱 함수가 다르기 때문에 버킷팅 파일 이름이 다르게 표시됩니다.

```
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-6-bucketing/day=01/
2020-10-21 20:35:16         13 000025_0
2020-10-21 20:35:22         14 000121_0
[hadoop@ip-10-0-0-122 ~]$ aws s3 ls s3://your-own-s3-bucket/emr-5-bucketing/day=01/
2020-10-21 20:32:07         13 000066_0
2020-10-21 20:32:51         14 000072_0
```

또한 EMR 6.x에서 Hive CLI를 통해 다음 명령을 실행하여 버전 차이를 확인할 수도 있습니다. 이 경우 버킷팅 버전 2가 반환됩니다.

```
hive> DESCRIBE FORMATTED test_bucketing;
...
Table Parameters:
    bucketing_version       2
...
```

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자](#)를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 5.31.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.31.0에 대한 정보가 포함됩니다. 변경 사항은 5.30.1에 관련됩니다.

최초 릴리스 날짜: 2020년 10월 9일

최종 업데이트 날짜: 2020년 10월 15일

업그레이드

- Amazon Glue 커넥터를 버전 1.13.0으로 업그레이드함
- Amazon SageMaker Spark SDK를 버전 1.4.0으로 업그레이드함
- Amazon Kinesis 커넥터를 버전 3.5.9로 업그레이드함
- 버전 1.11.852 AWS SDK for Java 로 업그레이드됨

- Bigtop-tomcat을 버전 8.5.56으로 업그레이드함
- EMR FS를 버전 2.43.0으로 업그레이드함
- EMR MetricsAndEventsApiGateway 클라이언트를 버전 1.4.0으로 업그레이드함
- EMR S3 Dist CP를 버전 2.15.0으로 업그레이드함
- EMR S3 Select를 버전 1.6.0으로 업그레이드함
- Flink를 버전 1.11.0으로 업그레이드함
- Hadoop을 버전 2.10.0으로 업그레이드함
- Hive를 버전 2.3.7로 업그레이드함
- Hudi를 버전 0.6.0으로 업그레이드함
- Hue를 버전 4.7.1로 업그레이드함
- JupyterHub가 버전 1.1.0으로 업그레이드됨
- Mxnet을 버전 1.6.0으로 업그레이드함
- OpenCV를 버전 4.3.0으로 업그레이드함
- Presto를 버전 0.238.3으로 업그레이드함
- TensorFlow를 버전 2.1.0으로 업그레이드함

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.
- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.

- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Hive 열 통계는 Amazon EMR 버전 5.31.0 이상에서 지원됩니다.
- 업그레이드된 구성 요소 버전.
- Amazon EMR 5.31.0에서 EMRFS S3EC V2 지원. S3 Java SDK 릴리스 1.11.837 이상에서 암호화 클라이언트 버전 2(S3EC V2)가 다양한 보안 개선 사항과 함께 도입되었습니다. 자세한 내용은 다음 자료를 참조하세요.
  - S3 블로그 게시물: Updates to the Amazon S3 encryption client.
  - AWS SDK for Java 개발자 안내서: 암호화 및 복호화 클라이언트를 V2로 마이그레이션합니다.
  - EMR 관리 안내서: Amazon S3 클라이언트 측 암호화

  암호화 클라이언트 V1은 이전 버전과의 호환성을 위해 SDK에서 계속 사용할 수 있습니다.

새로운 특성

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`
2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Amazon EMR 5.31.0에서는 Lake Formation과 통합되는 클러스터를 시작할 수 있습니다. 이 통합은 AWS Glue 데이터 카탈로그의 데이터베이스 및 테이블에 대한 세분화된 열 수준 데이터 필터링을 제공합니다. 또한 엔터프라이즈 자격 증명 시스템에서 EMR 노트북 또는 Apache Zeppelin에 연동

된 Single Sign-On을 사용할 수도 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon EMR과 AWS Lake Formation통합을 참조하세요.

Lake Formation을 사용하는 Amazon EMR은 현재 미국 동부(오하이오 및 버지니아 북부), 미국 서부(캘리포니아 북부 및 오레곤), 아시아 태평양(뭄바이, 서울, 싱가포르, 시드니 및 도쿄), 캐나다(중부), 유럽(프랑크푸르트, 아일랜드, 런던, 파리 및 스톡홀름), 남아메리카(상파울루)의 16개 AWS 리전에서 사용할 수 있습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

  > **Note**
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Amazon EMR 5.31.0 또는 5.32.0을 사용하는 클러스터에서 AtRestEncryption 또는 HDFS 암호화를 활성화하면 Hive 쿼리에서 다음과 같은 런타임 예외가 발생합니다.

```
TaskAttempt 3 failed, info=[Error: Error while running task ( failure ) :
 attempt_1604112648850_0001_1_01_000000_3:java.lang.RuntimeException:
 java.lang.RuntimeException: Hive Runtime Error while closing
 operators: java.io.IOException: java.util.ServiceConfigurationError:
 org.apache.hadoop.security.token.TokenIdentifier: Provider
 org.apache.hadoop.hbase.security.token.AuthenticationTokenIdentifier not found
```

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자](#)를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 6.1.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.1.0에 대한 정보가 포함됩니다. 변경 사항은 6.0.0에 관련됩니다.

최초 릴리스 날짜: 2020년 9월 4일

최종 업데이트 날짜: 2020년 10월 15일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.11.828
- Flink 버전 1.11.0
- Ganglia 버전 3.7.2
- Hadoop 버전 3.2.1-amzn-1

- HBase 버전 2.2.5
- HBase-operator-tools 1.0.0
- HCatalog 버전 3.1.2-amzn-0
- Hive 버전 3.1.2-amzn-1
- Hudi 버전 0.5.2-incubating
- Hue 버전 4.7.1
- JupyterHub 버전 1.1.0
- Livy 버전 0.7.0
- MXNet 버전 1.6.0
- Oozie 버전 5.2.0
- Phoenix 버전 5.0.0
- Presto 버전 0.232
- PrestoSQL 버전 338
- Spark 버전 3.0.0-amzn-0
- TensorFlow 버전 2.1.0
- Zeppelin 버전 0.9.0-preview1
- Zookeeper 버전 3.4.14
- 커넥터 및 드라이버: DynamoDB Connector 4.14.0

새로운 특성

- ARM 인스턴스 유형은 Amazon EMR 버전 5.30.0 및 Amazon EMR 버전 6.1.0부터 지원됩니다.
- M6g 범용 인스턴스 유형은 Amazon EMR 버전 6.1.0 및 5.30.0부터 지원됩니다. 자세한 내용은 Amazon EMR 관리 안내서에서 지원되는 인스턴스 유형을 참조하세요.
- EC2 배치 그룹 기능은 여러 프라이머리 노드 클러스터에 대한 옵션으로 Amazon EMR 버전 5.23.0부터 지원됩니다. 현재 배치 그룹 기능에서는 프라이머리 노드 유형만 지원되며 해당 프라이머리 노드에는 `SPREAD` 전략이 적용됩니다. `SPREAD` 전략은 하드웨어 장애 발생 시 여러 프라이머리 노드의 손실을 방지하기 위해 별도의 기본 하드웨어에 작은 그룹의 인스턴스를 배치합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR과 EC2 배치 그룹 통합을 참조하세요.
- Managed Scaling - Amazon EMR 버전 6.1.0에서는 Amazon EMR Managed Scaling을 활성화하여 워크로드에 따라 클러스터에서 인스턴스 또는 유닛 수를 자동으로 늘리거나 줄일 수 있습니다. Amazon EMR은 클러스터 지표를 지속적으로 평가하여 비용과 속도 측면에서 클러스터를 최적화하

는 조정 결정을 내립니다. Amazon EMR 버전 5.30.0 이상(6.0.0 제외)에서도 Managed Scaling을 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 리소스 조정을 참조하세요.

- PrestoSQL 버전 338은 EMR 6.1.0에서 지원됩니다. 자세한 내용은 Presto를 참조하세요.
  - PrestoSQL은 EMR 6.1.0 이상 버전에서만 지원되며, EMR 6.0.0 또는 EMR 5.x에서 지원되지 않습니다.
  - 애플리케이션 이름(`Presto`)은 클러스터에 PrestoDB를 설치하는 데 계속 사용됩니다. 클러스터에 PrestoSQL을 설치하려면 애플리케이션 이름, `PrestoSQL`을 사용합니다.
  - PrestoDB 또는 PrestoSQL 중 하나를 설치할 수 있지만 단일 클러스터에 둘 다 설치할 수는 없습니다. 클러스터를 생성하려고 할 때 PrestoDB와 PrestoSQL을 모두 지정하면 검증 오류가 발생하고 클러스터 생성 요청이 실패합니다.
  - PrestoSQL은 단일 마스터 클러스터 및 멀티 마스터 클러스터 모두에서 지원됩니다. 멀티 마스터 클러스터에서 PrestoSQL 또는 PrestoDB를 실행하려면 외부 Hive 메타스토어가 필요합니다. 여러 프라이머리 노드를 포함하는 EMR 클러스터에서 지원되는 애플리케이션을 참조하세요.
- Docker를 통해 Apache Hadoop 및 Apache Spark에서의 ECR 자동 인증 지원: Spark 사용자는 Docker Hub와 Amazon Elastic Container Registry(Amazon ECR)의 도커 이미지를 사용하여 환경 및 라이브러리 종속 항목을 정의할 수 있습니다.

  Docker를 구성하고 Amazon EMR 6.x를 사용하여 Docker와 함께 Spark 애플리케이션을 실행합니다.
- EMR에서 Apache Hive ACID 트랜잭션 지원: Amazon EMR 6.1.0에서는 Hive ACID 트랜잭션에 대한 지원을 추가하였고, 데이터베이스의 ACID 속성을 준수합니다. 이 기능을 사용하면 Amazon Simple Storage Service(S3)의 데이터를 사용하여 Hive 관리형 테이블에서 `INSERT, UPDATE, DELETE,` 및 `MERGE` 작업을 실행할 수 있습니다. 이는 스트리밍 수집, 데이터 복원, MERGE를 사용한 대량 업데이트, 느리게 변화하는 크기 등과 같은 사용 사례와 관련된 주요 기능입니다. 구성 예제 및 사용 사례를 비롯한 자세한 내용은 Amazon EMR supports Apache Hive ACID transactions를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했

습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.

- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Apache Flink는 EMR 6.0.0에서 지원되지 않지만 EMR 6.1.0에서 Flink 1.11.0으로 지원됩니다. 이 버전은 Hadoop 3을 공식적으로 지원하는 Flink의 첫 번째 버전입니다. Apache Flink 1.11.0 Release Announcement를 참조하세요.
- Ganglia는 기본 EMR 6.1.0 패키지 번들에서 제거되었습니다.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가

4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`
2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
```

```
done
sudo systemctl daemon-reload
```

- > **Important**
  >
  > Amazon EMR 6.1.0 및 6.2.0에는 모든 Hudi 삽입, 업서트 및 삭제 작업에 큰 영향을 미칠 수 있는 성능 문제가 포함되어 있습니다. Amazon EMR 6.1.0 또는 6.2.0에서 Hudi를 사용하려는 경우 AWS 지원팀에 문의하여 패치된 Hudi RPM을 받아야 합니다.

- `spark.driver.extraJavaOptions` 및 `spark.executor.extraJavaOptions`를 사용하여 사용자 지정 가비지 수집 구성을 설정하면 가비지 수집 구성 충돌로 인해 EMR 6.1에서 드라이버 및 실행기를 시작하지 못할 수 있습니다. EMR 릴리스 6.1.0에서는 `spark.driver.defaultJavaOptions` 및 `spark.executor.defaultJavaOptions` 속성 대신, 드라이버와 실행기에 대한 사용자 지정 Spark 가비지 수집 구성을 지정해야 합니다. 자세한 내용은 [Apache Spark Runtime Environment](#) 및 [Amazon EMR 6.1.0에서 Spark 가비지 수집 구성](#)을 참조하세요.
- Oozie에서 Pig를 사용하는 경우(또한 Hue는 Oozie 작업을 사용하여 Pig 스크립트를 실행하므로 Hue 내에서 Pig를 사용하는 경우) 네이티브 lzo 라이브러리를 로드할 수 없다는 오류가 발생합니다. 이 오류 메시지는 정보 제공용이며 Pig 실행을 차단하지 않습니다.
- Hudi 동시성 지원: 현재 Hudi는 단일 Hudi 테이블에 대한 동시 쓰기를 지원하지 않습니다. 또한 Hudi는 새 작성기를 시작하기 전에 진행 중인 작성기에서 수행하는 모든 변경 사항을 롤백합니다. 동시 쓰기는 이 메커니즘을 방해하고 경쟁 조건을 유발하여 데이터를 손상시킬 수 있습니다. 데이터 처리 워크플로의 일부로 항상 Hudi 테이블에서 단일 Hudi 작성기만 작동해야 합니다. Hudi는 동일한 Hudi 테이블에서 작동하는 여러 개의 동시 리더를 지원합니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Amazon EMR 6.1.0에는 Presto를 실행하는 클러스터에 영향을 미치는 문제가 있습니다. 오랜 기간(일 단위)이 경과한 후 클러스터에서 'su: failed to execute /bin/bash: Resource temporarily unavailable' 또는 'shell request failed on channel 0'과 같은 오류가 발생할 수 있습니다. 이 문제는 내부 Amazon EMR 프로세스(InstanceController)에서 너무 많은 경량 프로세스(LWP)를 생성하여 결국 Hadoop 사용자가 nproc 한도를 초과하기 때문에 발생합니다. 이를 통해 사용자는 추가 프로세스를 열 수 없습니다. 이 문제의 해결 방법은 EMR 6.2.0으로 업그레이드하는 것입니다.

릴리스 6.0.0

다음 릴리스 정보에는 Amazon EMR 릴리스 6.0.0에 대한 정보가 포함됩니다.

최초 릴리스 날짜: 2020년 3월 10일

지원되는 애플리케이션

- AWS SDK for Java 버전 1.11.711
- Ganglia 버전 3.7.2
- Hadoop 버전 3.2.1
- HBase 버전 2.2.3
- HCatalog 버전 3.1.2
- Hive 버전 3.1.2
- Hudi 버전 0.5.0-incubating
- Hue 버전 4.4.0
- JupyterHub 버전 1.0.0

- Livy 버전 0.6.0
- MXNet 버전 1.5.1
- Oozie 버전 5.1.0
- Phoenix 버전 5.0.0
- Presto 버전 0.230
- Spark 버전 2.4.4
- TensorFlow 버전 1.14.0
- Zeppelin 버전 0.9.0-SNAPSHOT
- Zookeeper 버전 3.4.14
- 커넥터 및 드라이버: DynamoDB Connector 4.14.0

> **Note**
>
> Flink, Sqoop, Pig 및 Mahout는 Amazon EMR 버전 6.0.0에서만 사용할 수 있습니다.

새로운 특성

- YARN 도커 런타임 지원 - Spark 작업과 같은 YARN 애플리케이션은 이제 도커 컨테이너의 컨텍스트에서 실행될 수 있습니다. 이를 통해 Amazon EMR 클러스터에 사용자 지정 라이브러리를 설치할 필요 없이 도커 이미지의 종속성을 쉽게 정의할 수 있습니다. 자세한 내용은 Docker 통합 구성 및 Amazon EMR 6.0.0을 사용하여 Docker와 함께 Spark 애플리케이션 실행을 참조하세요.
- Hive LLAP 지원 - Hive는 이제 향상된 쿼리 성능을 위해 LLAP 실행 모드를 지원합니다. 자세한 내용은 Hive LLAP 사용을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 Amazon EMR 조정에서 클러스터를 성공적으로 스케일 업 또는 스케일 다운하지 못하거나 애플리케이션 장애가 발생한 경우 이 기능과 관련된 문제를 수정하기 위한 릴리스입니다.
- Amazon EMR 클러스터 내 대몬(daemon)이 YARN 노드 상태 및 HDFS 노드 상태 수집과 같은 상태 확인 활동을 실행할 때 활용도가 높은 대규모 클러스터에 대한 조정 요청이 실패하는 문제를 수정했습니다. 이는 클러스터 내 대몬(daemon)이 내부 Amazon EMR 구성 요소로 노드의 상태 데이터를 전달할 수 없었기 때문에 발생했습니다.

- 조정 작업 중에 신뢰성을 개선하기 위해 IP 주소를 재사용할 때 노드 상태를 올바르게 추적하도록 EMR 클러스터 내 대몬(daemon)을 개선했습니다.
- SPARK-29683. Spark에서는 사용 가능한 모든 노드가 거부 목록에 있다고 가정했기 때문에 클러스터 스케일 다운 중에 작업 실패가 발생하는 문제를 수정했습니다.
- YARN-9011. 클러스터를 스케일 업하거나 스케일 다운하려고 할 때 YARN 서비스 해제의 경쟁 조건으로 인해 작업 실패가 발생하는 문제를 수정했습니다.
- Amazon EMR 클러스터 내 대몬(daemon)과 YARN 및 HDFS 사이에서 노드 상태가 항상 일관되도록 보장함으로써 클러스터 조정 중 단계 또는 작업이 실패하는 문제를 수정했습니다.
- Kerberos 인증으로 활성화된 Amazon EMR 클러스터에서 스케일 다운 및 단계 제출과 같은 클러스터 작업이 실패하는 문제를 수정했습니다. 이는 Amazon EMR 클러스터 내 대몬(daemon)이 프라이머리 노드에서 실행되는 HDFS 및 YARN과 안전하게 통신하는 데 필요한 Kerberos 티켓을 갱신하지 않았기 때문입니다.
- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- Amazon Linux
  - Amazon Linux 2는 EMR 6.x 릴리스 시리즈의 운영 체제입니다.
  - `systemd`는 Amazon Linux 1에서 사용되는 `upstart` 대신 서비스 관리에 사용됩니다.
- Java Development Kit(JDK)
  - Coretto JDK 8은 EMR 6.x 릴리스 시리즈의 기본 JDK입니다.
- Scala
  - Scala 2.12는 Apache Spark 및 Apache Livy와 함께 사용됩니다.
- Python 3
  - Python 3은 이제 EMR에서 Python의 기본 버전입니다.
- YAR의 노드 레이블
  - Amazon EMR 6.x 릴리스 시리즈부터 YARN 노드 레이블 기능은 기본적으로 비활성화되어 있습니다. 애플리케이션 마스터 프로세스는 기본적으로 코어 및 작업 노드 모두에서 실행할 수 있습니다. `yarn.node-labels.enabled` 및 `yarn.node-labels.am.default-node-label-expression` 속성을 구성하여 YARN 노드 레이블 기능을 활성화할 수 있습니다. 자세한 내용은 프라이머리, 코어 및 태스크 노드 이해를 참조하세요.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

     `LimitNOFILE=65536`

     `LimitNPROC=65536`

  2. InstanceController 다시 시작

     `$ sudo systemctl daemon-reload`

     `$ sudo systemctl restart instance-controller`

  부트스트랩 작업(BA)을 사용하여 ulimit 설정

  클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
```

```
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- PySpark, SparkR 및 spark-shell을 포함한 Spark 대화형 셸은 도커를 추가 라이브러리와 함께 사용하는 것을 지원하지 않습니다.
- Amazon EMR 버전 6.0.0에서 Python 3을 사용하려면 `yarn.nodemanager.env-whitelist`에 PATH를 추가해야 합니다.
- Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용하는 경우 LLAP(Live Long and Process) 기능은 지원되지 않습니다.
- Spark 및 Docker 통합과 함께 Amazon EMR 6.0.0을 사용하는 경우 Docker 런타임으로 Spark 작업을 제출할 때 실패를 방지하기 위해 클러스터의 인스턴스를 동일한 인스턴스 유형과 동일한 크기의 EBS 볼륨으로 구성해야 합니다.
- Amazon EMR 6.0.0의 경우 Amazon S3 스토리지 모드에서 HBase는 HBASE-24286 문제의 영향을 받습니다. 기존 S3 데이터를 사용하여 클러스터를 생성할 때는 HBase 마스터를 초기화할 수 없습니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.

- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.30.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.30.1에 대한 정보가 포함됩니다. 변경 사항은 5.30.0에 관련됩니다.

최초 릴리스 날짜: 2020년 6월 30일

최종 업데이트 날짜: 2020년 8월 24일

변경 사항, 향상된 기능 및 해결된 문제

- 최신 Amazon EMR 릴리스에서는 Amazon EMR의 이전 AL2에서 더 낮은 '최대 열린 파일' 한도와 관련된 문제가 수정되었습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 이제 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다.
- 인스턴스 컨트롤러 프로세스에서 무한한 수의 프로세스를 생성하는 문제가 수정되었습니다.
- Hue에서 Hive 쿼리를 실행할 수 없어서 'database is locked' 메시지를 표시하고 쿼리가 실행되지 않는 문제가 수정되었습니다.
- EMR 클러스터에서 더 많은 작업을 동시에 실행할 수 있도록 Spark 문제가 수정되었습니다.
- Jupyter 서버에서 'too many files open error'를 발생시키는 Jupyter Notebook 문제가 수정되었습니다.
- 클러스터 시작 시간 관련 문제가 수정되었습니다.

새로운 특성

- Amazon EMR 버전 6.x 및 EMR 버전 5.30.1 이상에서 Tez UI 및 YARN 타임라인 서버의 영구 애플리케이션 인터페이스를 사용할 수 있습니다. 영구 애플리케이션 기록에 대한 원클릭 링크 액세스를 사용하면 SSH 연결을 통해 웹 프록시를 설정하지 않고도 작업 기록에 빠르게 액세스할 수 있습니다. 활성 클러스터와 종료된 클러스터에 대한 로그는 애플리케이션 종료 후 30일 동안 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 영구 애플리케이션 사용자 인터페이스 보기를 참조하세요.
- EMR 노트북 실행 API는 스크립트 또는 명령줄을 통해 EMR 노트북을 실행하는 데 사용할 수 있습니다. AWS 콘솔 없이 EMR 노트북 실행을 시작, 중지, 나열 및 설명할 수 있으므로 EMR 노트북을 프로그래밍 방식으로 제어할 수 있습니다. 파라미터화된 노트북 셀을 사용하면 새로운 파라미터 값 세트마다 노트북의 복사본을 만들지 않고도 노트북에 다양한 파라미터 값을 전달할 수 있습니다. EMR API 작업을 참조하세요. 샘플 코드는 EMR Notebooks를 프로그래밍 방식으로 실행하는 샘플 명령을 참조하세요.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

  이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

  명령줄에서 ulimit를 명시적으로 설정합니다.

  1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

     ```
     LimitNOFILE=65536
     ```

```
LimitNPROC=65536
```

2. InstanceController 다시 시작

```
$ sudo systemctl daemon-reload
```

```
$ sudo systemctl restart instance-controller
```

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- EMR Notebooks

  클러스터 프라이머리 노드에 커널과 추가 Python 라이브러리를 설치할 수 있는 기능은 EMR 버전 5.30.1에서 기본적으로 비활성화됩니다. 이 기능에 대한 자세한 내용은 클러스터 프라이머리 노드에 커널 및 Python 라이브러리 설치를 참조하세요.

  이 기능을 활성화하려면 다음을 수행합니다.

  1. EMR Notebooks의 서비스 역할에 연결된 권한 정책이 다음 작업을 허용하는지 확인합니다.

```
elasticmapreduce:ListSteps
```

자세한 내용은 [EMR 노트북 서비스 역할] 단원을 참조하십시오.

2. AWS CLI 를 사용하여 다음 예제와 같이 EMR 노트북을 설정하는 클러스터에서 단계를 실행합니다. *us-east-1*을 클러스터가 상주하는 리전으로 바꿉니다. 자세한 내용은 [AWS CLI를 사용하여 클러스터에 단계 추가]를 참조하세요.

```
aws emr add-steps  --cluster-id MyClusterID --steps
 Type=CUSTOM_JAR,Name=EMRNotebooksSetup,ActionOnFailure=CONTINUE,Jar=s3://us-
east-1.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://
awssupportdatasvcs.com/bootstrap-actions/EMRNotebooksSetup/emr-notebooks-setup.sh"]
```

- Managed Scaling

  Presto가 설치되지 않은 5.30.0 및 5.30.1 클러스터에서 Managed Scaling을 수행하면 애플리케이션 장애가 발생하거나 균일한 인스턴스 그룹 또는 인스턴스 플릿이 ARRESTED 상태로 유지될 수 있습니다. 특히 스케일 다운 작업 이후 바로 스케일 업 작업이 수행되는 경우가 이에 해당합니다.

  해결 방법으로 작업에 Presto가 필요하지 않더라도 Amazon EMR 릴리스 5.30.0 및 5.30.1에서 클러스터를 생성할 때 설치할 애플리케이션으로 Presto를 선택합니다.

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

  ```
  kinit -kt <keytab_file> <principal>
  ```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.
  - 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
  - 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
  - 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 [UTF-8 인코딩 테이블 및 유니코드 문자](#)를 참조하세요.

  이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 5.30.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.30.0에 대한 정보가 포함됩니다. 변경 사항은 5.29.0에 관련됩니다.

최초 릴리스 날짜: 2020년 5월 13일

최종 업데이트 날짜: 2020년 6월 25일

업그레이드

- 버전 1.11.759 AWS SDK for Java 로 업그레이드됨
- Amazon SageMaker Spark SDK가 버전 1.3.0으로 업그레이드됨
- EMR 레코드 서버가 버전 1.6.0으로 업그레이드됨
- Flink가 버전 1.10.0으로 업그레이드됨

- Ganglia가 버전 3.7.2로 업그레이드됨
- HBase가 버전 1.4.13으로 업그레이드됨
- Hudi가 버전 0.5.2-incubating으로 업그레이드됨
- Hue가 버전 4.6.0으로 업그레이드됨
- JupyterHub가 버전 1.1.0으로 업그레이드됨
- Livy가 버전 0.7.0-incubating으로 업그레이드됨
- Oozie가 버전 5.2.0으로 업그레이드됨
- Presto가 버전 0.232로 업그레이드됨
- Spark가 버전 2.4.5로 업그레이드됨
- 업그레이드된 커넥터 및 드라이버: Amazon Glue 커넥터 1.12.0, Amazon Kinesis 커넥터 3.5.0, EMR DynamoDB 커넥터 4.14.0

새로운 특성

- EMR Notebooks - 5.30.0을 사용하여 생성된 EMR 클러스터와 함께 사용하면 EMR 노트북 커널이 클러스터에서 실행됩니다. 이로써 노트북 성능이 향상되며, 커널을 설치하고 사용자 지정할 수 있습니다. 클러스터 프라이머리 노드에 Python 라이브러리를 설치할 수도 있습니다. 자세한 내용은 EMR 관리 안내서에서 커널 및 라이브러리 설치 및 사용을 참조하세요.
- Managed Scaling - Amazon EMR 버전 5.30.0 이상에서는 EMR Managed Scaling을 활성화하여 워크로드에 따라 클러스터에서 인스턴스 또는 유닛 수를 자동으로 늘리거나 줄일 수 있습니다. Amazon EMR은 클러스터 지표를 지속적으로 평가하여 비용과 속도 측면에서 클러스터를 최적화하는 조정 결정을 내립니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 리소스 조정을 참조하세요.
- Amazon S3에 저장된 로그 파일 암호화 - Amazon EMR 버전 5.30.0 이상에서는 AWS KMS 고객 관리형 키를 사용하여 Amazon S3에 저장된 로그 파일을 암호화할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon S3에 저장된 로그 파일 암호화를 참조하세요.
- Amazon Linux 2 지원 - EMR 버전 5.30.0 이상에서 EMR은 Amazon Linux 2 OS를 사용합니다. 새로운 사용자 지정 Amazon Machine Image(AMI)는 Amazon Linux 2 AMI를 기반으로 해야 합니다. 자세한 내용은 사용자 지정 AMI 사용을 참조하십시오.
- Presto 단계적 자동 조정 - 5.30.0 버전을 사용하는 EMR 클러스터의 경우 Auto Scaling 제한 시간을 설정할 수 있습니다. 이 제한 시간을 통해 Presto에서 노드가 서비스 해제되기 전에 실행을 완료할 수 있습니다. 자세한 내용은 단계적 서비스 해제 기능이 있는 자동 조정 사용 단원을 참조하십시오.
- 새 할당 전략 옵션을 사용한 플릿 인스턴스 생성 - 새 할당 전략 옵션은 EMR 버전 5.12.1 이상에서 사용할 수 있습니다. 이를 통해 클러스터 프로비저닝 속도가 빨라지고 스팟 할당이 더 정확해지며 스

팟 인스턴스 중단이 줄어듭니다. 기본이 아닌 EMR 서비스 역할은 업데이트해야 합니다. 인스턴스 플릿 구성을 참조하세요.

- sudo systemctl stop 및 sudo systemctl start 명령 – Amazon Linux 2 OS를 사용하는 EMR 버전 5.30.0 이상에서 EMR은 `sudo systemctl stop` 및 `sudo systemctl start` 명령을 사용하여 서비스를 다시 시작합니다. 자세한 내용은 Amazon EMR에서 서비스를 다시 시작하려면 어떻게 해야 하나요?를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- EMR 버전 5.30.0은 기본적으로 Ganglia를 설치하지 않습니다. 클러스터 생성 시 Ganglia를 설치하도록 명시적으로 선택할 수 있습니다.
- Spark 성능 최적화
- Presto 성능 최적화
- Python 3은 Amazon EMR 버전 5.30.0 이상의 기본값입니다.
- 프라이빗 서브넷의 서비스 액세스에 대한 기본 관리형 보안 그룹이 새 규칙으로 업데이트되었습니다. 서비스 액세스에 대한 사용자 지정 보안 그룹을 사용하는 경우, 기본 관리형 보안 그룹과 동일한 규칙을 포함시켜야 합니다. 자세한 내용은 서비스 액세스에 대한 Amazon EMR 관리형 보안 그룹(프라이빗 서브넷)을 참조하세요. Amazon EMR에 대해 사용자 지정 서비스 역할을 사용하는 경우 EMR에서 보안 그룹이 올바르게 생성되는지 여부를 검증할 수 있도록 `ec2:describeSecurityGroups`에 대한 권한을 부여해야 합니다. `EMR_DefaultRole`을 사용하는 경우 이 권한은 이미 기본 관리형 정책에 포함되어 있습니다.

알려진 문제

- 이전 AL2에서 더 낮은 '최대 열린 파일' 한도[최신 릴리스에서 수정됨]. Amazon EMR 릴리스: emr-5.30.x, emr-5.31.0, emr-5.32.0, emr-6.0.0, emr-6.1.0, emr-6.2.0은 Amazon Linux 2(AL2)의 이전 버전에 기반합니다. 이전 버전에서는 Amazon EMR 클러스터가 기본 AMI에서 생성되는 경우 '최대 열린 파일'에 대한 ulimit 설정이 더 낮습니다. Amazon EMR 릴리스 5.30.1, 5.30.2, 5.31.1, 5.32.1, 6.0.1, 6.1.1, 6.2.1, 5.33.0, 6.3.0 이상에는 더 높은 '최대 열린 파일' 설정을 포함한 영구적 수정 사항이 포함되어 있습니다. 열린 파일 한도가 낮은 릴리스의 경우 Spark 작업을 제출할 때 'Too many open files' 오류가 발생합니다. 영향을 받는 릴리스에서 Amazon EMR 기본 AMI의 '최대 열린 파일'에 대한 기본 ulimit 설정은 4,096개이며, 이는 최신 Amazon Linux 2 AMI의 65,536개 파일 제한보다 낮습니다. '최대 열린 파일'에 대한 더 낮은 ulimit 설정을 사용하면 Spark 드라이버와 실행기가 4,096개가 넘는 파일을 열려고 할 때 Spark 작업이 실패합니다. 이 문제를 해결하기 위해 Amazon EMR에는 클러스터 생성 시 ulimit 설정을 조정하는 부트스트랩 작업(BA) 스크립트가 있습니다.

이 문제에 대한 영구 수정 사항이 포함되지 않은 이전 Amazon EMR 버전을 사용하는 경우 다음 해결 방법을 통해 instance-controller ulimit를 최대 65,536개의 파일로 명시적으로 설정할 수 있습니다.

명령줄에서 ulimit를 명시적으로 설정합니다.

1. `/etc/systemd/system/instance-controller.service`를 편집하여 Service 섹션에 다음 파라미터를 추가합니다.

   `LimitNOFILE=65536`

   `LimitNPROC=65536`

2. InstanceController 다시 시작

   `$ sudo systemctl daemon-reload`

   `$ sudo systemctl restart instance-controller`

부트스트랩 작업(BA)을 사용하여 ulimit 설정

클러스터 생성 시 부트스트랩 작업(BA) 스크립트를 사용하여 instance-controller ulimit를 65,536개 파일로 구성할 수도 있습니다.

```
#!/bin/bash
for user in hadoop spark hive; do
sudo tee /etc/security/limits.d/$user.conf << EOF
$user - nofile 65536
$user - nproc 65536
EOF
done
for proc in instancecontroller logpusher; do
sudo mkdir -p /etc/systemd/system/$proc.service.d/
sudo tee /etc/systemd/system/$proc.service.d/override.conf << EOF
[Service]
LimitNOFILE=65536
LimitNPROC=65536
EOF
pid=$(pgrep -f aws157.$proc.Main)
sudo prlimit --pid $pid --nofile=65535:65535 --nproc=65535:65535
done
sudo systemctl daemon-reload
```

- Managed Scaling

  Presto가 설치되지 않은 5.30.0 및 5.30.1 클러스터에서 Managed Scaling을 수행하면 애플리케이션 장애가 발생하거나 균일한 인스턴스 그룹 또는 인스턴스 플릿이 ARRESTED 상태로 유지될 수 있습니다. 특히 스케일 다운 작업 이후 바로 스케일 업 작업이 수행되는 경우가 이에 해당합니다.

  해결 방법으로 작업에 Presto가 필요하지 않더라도 Amazon EMR 릴리스 5.30.0 및 5.30.1에서 클러스터를 생성할 때 설치할 애플리케이션으로 Presto를 선택합니다.

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 /etc/hadoop.keytab에 있으며, 보안 주체는 hadoop/<hostname>@<REALM> 양식입니다.

  > **Note**
  >
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

- Hue 4.6.0의 기본 데이터베이스 엔진은 SQLite이며, 이 때문에 외부 데이터베이스에서 Hue를 사용하려고 하면 문제가 발생합니다. 이 문제를 해결하려면 hue-ini 구성 분류에서 engine을 mysql로 설정합니다. 이 문제는 Amazon EMR 버전 5.30.1에서 수정되었습니다.
- Hive 파티션 위치 형식과 함께 Spark를 사용하여 Amazon S3의 데이터를 읽고, Amazon EMR 릴리스 5.30.0~5.36.0 및 6.2.0~6.9.0에서 Spark를 실행하는 경우 클러스터가 데이터를 제대로 읽지 못하

는 문제가 발생할 수 있습니다. 파티션에 다음과 같은 특성이 모두 있는 경우 이러한 상황이 발생할 수 있습니다.

- 둘 이상의 파티션이 동일한 테이블에서 스캔됩니다.
- 하나 이상의 파티션 디렉터리 경로가 하나 이상의 다른 파티션 디렉터리 경로의 접두사입니다. 예를 들어 `s3://bucket/table/p=a`는 `s3://bucket/table/p=a b`의 접두사입니다.
- 다른 파티션 디렉터리에서 접두사 뒤에 오는 첫 번째 문자의 UTF-8 값이 / 문자(U+002F)보다 작습니다. 예를 들어 `s3://bucket/table/p=a b`에서 a와 b 사이에 있는 공백 문자(U+0020)가 이 카테고리에 속합니다. `!"#$%&‘()*+,-`와 같은 14개의 비제어 문자가 있습니다. 자세한 내용은 UTF-8 인코딩 테이블 및 유니코드 문자를 참조하세요.

이 문제를 해결하려면 `spark-defaults` 분류에서 `spark.sql.sources.fastS3PartitionDiscovery.enabled` 구성을 `false`로 설정합니다.

릴리스 5.29.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.29.0에 대한 정보가 포함됩니다. 변경 사항은 5.28.1에 관련됩니다.

최초 릴리스 날짜: 2020년 1월 17일

업그레이드

- 버전 1.11.682 AWS SDK for Java 로 업그레이드됨
- Hive가 버전 2.3.6으로 업그레이드됨
- Flink가 버전 1.9.1로 업그레이드됨
- EmrFS가 버전 2.38.0으로 업그레이드됨
- EMR DynamoDB Connector가 버전 4.13.0으로 업그레이드됨

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Spark 성능 최적화
- EMRFS
  - 일관된 보기를 위해 관리 안내서에서 emrfs-site.xml 기본 설정이 업데이트됩니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.28.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.28.1에 대한 정보가 포함됩니다. 변경 사항은 5.28.0에 관련됩니다.

최초 릴리스 날짜: 2020년 1월 10일

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Spark 호환성 문제를 수정했습니다.

- CloudWatch 지표
  - 여러 프라이머리 노드가 있는 EMR 클러스터에 Amazon CloudWatch 지표가 게시되는 문제를 수정했습니다.
- 로그 메시지 비활성화
  - false 로그 메시지 "이전 버전(<4.5.8)의 Apache http 클라이언트를 사용..."을 비활성화했습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:
  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 `hadoop` 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 `hadoop` 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

  > **Note**
  >
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.28.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.28.0에 대한 정보가 포함됩니다. 변경 사항은 5.27.0에 관련됩니다.

최초 릴리스 날짜: 2019년 11월 12일

업그레이드

- Flink가 버전 1.9.0으로 업그레이드됨
- Hive가 버전 2.3.6으로 업그레이드됨
- MXNet이 1.5.1로 업그레이드됨
- Phoenix가 버전 4.14.3으로 업그레이드됨
- Presto가 버전 0.227로 업그레이드됨
- Zeppelin이 버전 0.8.2로 업그레이드됨

새로운 특성

- 이제 Apache Hudi는 클러스터를 생성할 때 Amazon EMR에서 설치할 수 있습니다. 자세한 내용은 Hudi 단원을 참조하십시오.
- (2019년 11월 25일) 이제 클러스터 사용률을 개선하고 비용을 절감하기 위해 여러 단계를 병렬로 실행하도록 선택할 수 있습니다. 대기 중인 단계와 실행 중인 단계를 모두 취소할 수도 있습니다. 자세한 내용은 AWS CLI 및 콘솔을 사용하여 단계 작업을 참조하세요.
- (2019년 12월 3일) 이제 온프레미스 시설에서 네이티브 AWS 서비스, 인프라 및 운영 모델을 AWS Outposts AWS Outposts 활성화하여 EMR 클러스터를 생성하고 실행할 수 있습니다. AWS Outposts 환경에서는 AWS 클라우드에서 사용하는 것과 동일한 AWS APIs, 도구 및 인프라를 사용할 수 있습니다. 자세한 내용은 의 EMR 클러스터를 참조하세요 AWS Outposts.
- (2020년 3월 11일) Amazon EMR 버전 5.28.0부터 AWS 로컬 영역을 지원하는 AWS 리전의 논리적 확장으로 로컬 영역 서브넷에서 Amazon EMR 클러스터를 생성하고 실행할 수 있습니다. 로컬 영역을 사용하면 Amazon EMR 기능과 컴퓨팅 및 스토리지 AWS 서비스와 같은 서비스 하위 집합을 사용자와 더 가까운 곳에 배치하여 로컬에서 실행되는 애플리케이션에 대한 매우 짧은 지연 시간 액세스를 제공할 수 있습니다. 사용 가능한 로컬 영역 목록은 AWS Local Zones를 참조하세요. 사용 가능한 AWS 로컬 영역에 액세스하는 방법에 대한 자세한 내용은 리전, 가용 영역 및 로컬 영역을 참조하세요.

  로컬 영역은 현재 Amazon EMR Notebooks를 지원하지 않으며 인터페이스 VPC 엔드포인트(AWS PrivateLink)를 사용하여 Amazon EMR에 직접 연결하는 기능도 지원하지 않습니다.

변경 사항, 향상된 기능 및 해결된 문제

- 고가용성 클러스터에 대해 확장된 애플리케이션 지원

  - 자세한 내용은 Amazon EMR 관리 안내서에서 [여러 프라이머리 노드가 있는 EMR 클러스터에서 지원되는 애플리케이션]을 참조하십시오.
- Spark
  - 성능 최적화
- Hive
  - 성능 최적화
- Presto
  - 성능 최적화

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

  > **Note**
  >
  > 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.27.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.27.0에 대한 정보가 포함됩니다. 변경 사항은 5.26.0에 관련됩니다.

최초 릴리스 날짜: 2019년 9월 23일

업그레이드

- AWS SDK for Java 1.11.615
- Flink 1.8.1
- JupyterHub 1.0.0
- Spark 2.4.4
- TensorFlow 1.14.0
- 커넥터 및 드라이버:
  - DynamoDB 커넥터 4.12.0

새로운 특성

- (2019년 10월 24일) 다음과 같은 EMR 노트북의 새로운 기능은 모든 Amazon EMR 릴리스에서 사용할 수 있습니다.
  - 이제 Git 리포지토리를 EMR 노트북과 연결하여 버전 제어 환경에서 노트북을 저장할 수 있습니다. 원격 Git 리포지토리를 통해 동료와 코드를 공유할 수 있고 기존 Jupyter Notebook을 재사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon EMR Notebooks에 Git 리포지토리 연결을 참조하세요.
  - 이제 EMR 노트북에서 nbdime 유틸리티를 사용하여 노트북 비교 및 병합을 간소화할 수 있습니다.
  - 이제 EMR 노트북이 JupyterLab을 지원합니다. JupyterLab은 Jupyter Notebook과 완전히 호환 가능한 웹 기반 대화식 개발 환경입니다. 이제 JupyterLab 또는 Jupyter Notebook 편집기에서 노트북을 열 수 있는 옵션이 제공됩니다.
- (2019년 10월 30일) Amazon EMR 버전 5.25.0 이상부터 콘솔에서 클러스터 요약 페이지 또는 애플리케이션 기록 탭을 사용해 Spark 기록 서버 UI에 연결할 수 있습니다. SSH 연결을 통해 웹 프록시를 설정하는 대신, Spark 기록 서버 UI에 액세스하여 애플리케이션 지표를 보고 활성 클러스터 및 종료된 클러스터에 대한 관련 로그 파일에 액세스할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 영구 애플리케이션 사용자 인터페이스에 대해 클러스터 외부에서 액세스를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터
  - 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에 Flink를 설치하고 실행할 수 있습니다. 자세한 내용은 지원되는 애플리케이션 및 기능을 참조하세요.
  - 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에서 HDFS 투명 암호화를 구성할 수 있습니다. 자세한 내용은 여러 프라이머리 노드를 포함하는 EMR 클러스터에서 HDFS의 투명한 암호화를 참조하세요.
  - 이제 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에서 실행 중인 애플리케이션의 구성을 수정할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.
- Amazon EMR-DynamoDB 커넥터
  - 이제 Amazon EMR-DynamoDB 커넥터에서 부울, 목록, 맵, 항목, null과 같은 DynamoDB 데이터 유형을 지원합니다. 자세한 정보는 Hive 테이블을 설정하여 Hive 명령 실행을 참조하십시오.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

Release 5.26.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.26.0에 대한 정보가 포함됩니다. 변경 사항은 5.25.0에 관련됩니다.

최초 릴리스 날짜: 2019년 8월 8일

최종 업데이트 날짜: 2019년 8월 19일

업그레이드

- AWS SDK for Java 1.11.595
- HBase 1.4.10
- Phoenix 4.14.2
- 커넥터 및 드라이버:
  - DynamoDB 커넥터 4.11.0
  - MariaDB 커넥터 2.4.2
  - Amazon Redshift JDBC 드라이버 1.2.32.1056

새로운 특성

- (베타) Amazon EMR 5.26.0에서 Lake Formation과 통합된 클러스터를 시작할 수 있습니다. 이 통합은 AWS Glue 데이터 카탈로그의 데이터베이스 및 테이블에 대한 세분화된 열 수준 액세스를 제공합니다. 또한 엔터프라이즈 자격 증명 시스템에서 EMR 노트북 또는 Apache Zeppelin에 연동된 Single Sign-On을 사용할 수도 있습니다. 자세한 내용은 [Integrating Amazon EMR with AWS Lake Formation (Beta)]를 참조하세요.
- (2019년 8월 19일) 보안 그룹을 지원하는 모든 Amazon EMR 릴리스에 Amazon EMR 퍼블릭 액세스 차단 기능을 사용할 수 있습니다. 퍼블릭 액세스 차단은 각 AWS 리전에 적용되는 계정 전체의 설정입니다. 포트가 예외로 지정되지 않은 한, 퍼블릭 액세스 차단은 클러스터와 연결된 보안 그룹에

포트의 IPv4 0.0.0.0/0 또는 IPv6 ::/0(퍼블릭 액세스)에서 인바운드 트래픽을 허용하는 규칙이 있을 때 클러스터가 시작되지 않도록 합니다. 포트 22는 기본적으로 예외로 설정됩니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon EMR 퍼블랙 액세스 차단 기능 사용을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- EMR Notebooks
  - EMR 5.26.0 이상에서 EMR Notebooks는 기본 Python 라이브러리 외에도 노트북 범위의 Python 라이브러리를 지원합니다. 클러스터를 다시 생성하거나 노트북을 클러스터에 다시 연결할 필요 없이 노트북 편집기 내에서 노트북 범위 라이브러리를 설치할 수 있습니다. 노트북 범위의 라이브러리는 Python 가상 환경에서 생성되므로 현재 노트북 세션에만 적용됩니다. 이를 통해 노트북 종속성을 분리할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 노트북 범위 라이브러리 사용을 참조하세요.
- EMRFS
  - `fs.s3.consistent.metadata.etag.verification.enabled`를 `true`로 설정하여 ETag 확인 기능(베타)을 활성화할 수 있습니다. 이 기능을 통해 EMRFS는 Amazon S3 ETag를 사용하여 읽고 있는 객체가 사용 가능한 최신 버전인지 확인합니다. 이 기능은 동일한 이름을 유지하면서 Amazon S3의 파일을 덮어쓰는 업데이트 후 읽기 사용 사례에 유용합니다. 이 ETag 확인 기능은 현재 S3 Select에서 작동하지 않습니다. 자세한 내용은 일관된 보기 구성 단원을 참조하십시오.
- Spark
  - 동적 파티션 잘라내기, INTERSECT에 앞선 DISTINCT, JOIN-DISINCT 쿼리에 대한 SQL 계획 통계 추론의 개선, 스칼라 하위 쿼리 평면화, 최적화된 조인 재정렬 및 블룸 필터 조인과 같은 최적화가 이제 기본적으로 설정됩니다. 자세한 내용은 Spark 성능 최적화를 참조하십시오.
  - 정렬 병합 조인의 전체 스테이지 코드 생성 기능이 개선되었습니다.
  - 쿼리 조각 및 하위 쿼리 재사용이 개선되었습니다.
  - Spark 시작 시 실행기를 사전 할당하도록 개선되었습니다.
  - 조인의 작은 쪽에 브로드캐스트 힌트가 포함되어 있으면 블룸 필터 조인이 더 이상 적용되지 않습니다.
- Tez의 Hive 실행 시간 비교
  - Tez 관련 문제를 해결했습니다. 이제 Tez UI는 여러 프라이머리 노드를 포함하는 Amazon EMR 클러스터에서 작동합니다.

알려진 문제

- 정렬 병합 조인에 대한 향상된 전체 스테이지 코드 생성 기능을 사용하면 메모리 압력을 높일 수 있습니다. 이 최적화는 성능을 향상시키지만 `spark.yarn.executor.memoryOverheadFactor`가 충분한 메모리를 제공하도록 조정되지 않은 경우 작업 재시도 또는 실패를 초래할 수 있습니다. 이 기능을 비활성화하려면 `spark.sql.sortMergeJoinExec.extendedCodegen.enabled`를 false로 설정하십시오.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

    ```
    kinit -kt <keytab_file> <principal>
    ```

    일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.25.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.25.0에 대한 정보가 포함됩니다. 변경 사항은 5.24.1에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 17일

최종 업데이트 날짜: 2019년 10월 30일

Amazon EMR 5.25.0

업그레이드

- AWS SDK for Java 1.11.566
- Hive 2.3.5
- Presto 0.220
- Spark 2.4.3
- TensorFlow 1.13.1
- Tez 0.9.2
- Zookeeper 3.4.14

새로운 특성

- (2019년 10월 30일) Amazon EMR 버전 5.25.0부터 콘솔에서 클러스터 요약 페이지 또는 애플리케이션 기록 탭을 사용해 Spark 기록 서버 UI에 연결할 수 있습니다. SSH 연결을 통해 웹 프록시를 설정하는 대신, Spark 기록 서버 UI에 액세스하여 애플리케이션 지표를 보고 활성 클러스터 및 종료된 클러스터에 대한 관련 로그 파일에 액세스할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 영구 애플리케이션 사용자 인터페이스에 대해 클러스터 외부에서 액세스를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Bloom 필터로 입력을 사전 필터링하여 일부 조인의 성능을 개선했습니다. 이 최적화는 기본적으로 비활성화되어 있으며, Spark 구성 파라미터 `spark.sql.bloomFilterJoin.enabled`를 `true`로 설정하여 활성화할 수 있습니다.
  - 문자열 형식 열을 기준으로 그룹화하는 성능을 개선했습니다.
  - HBase를 설치하지 않고 클러스터에 대한 R4 인스턴스 유형의 핵심 구성 및 기본 Spark 실행기 메모리를 개선했습니다.
  - 잘린 테이블이 조인의 왼쪽에 있어야 하는 이전의 동적 파티션 잘라내기 기능에 대한 문제를 해결했습니다.
  - 별칭과 관련된 추가 사례를 적용하기 위해 INTERSECT 이전 DISTINCT 최적화를 개선했습니다.

  - DISTINCT 쿼리 이전 JOIN에 대한 SQL 계획 통계 추론을 개선했습니다. 이 개선 사항은 기본적으로 비활성화되어 있으며, Spark 구성 파라미터 `spark.sql.statsImprovements.enabled`를 `true`로 설정하여 활성화할 수 있습니다. 이 최적화는 Intersect 이전 Distinct 기능에 필요하며 `spark.sql.optimizer.distinctBeforeIntersect.enabled`를 `true`로 설정할 경우 자동으로 활성화됩니다.
  - 테이블 크기 및 필터에 따라 조인 순서를 최적화했습니다. 이 최적화는 기본적으로 비활성화되어 있으며, Spark 구성 파라미터 `spark.sql.optimizer.sizeBasedJoinReorder.enabled`를 `true`로 설정하여 활성화할 수 있습니다.

  자세한 내용은 Spark 성능 최적화를 참조하십시오.
- EMRFS
  - 이제 EMRFS 설정 `fs.s3.buckets.create.enabled`가 기본적으로 비활성화됩니다. 테스트 결과 이 설정을 비활성화하면 성능이 향상되고 의도치 않은 S3 버킷 생성이 방지된다는 사실을 확인했습니다. 애플리케이션이 이 기능을 사용하는 경우 `emrfs-site` 구성 분류에서 속성 `fs.s3.buckets.create.enabled`를 `true`로 설정하여 활성화할 수 있습니다. 자세한 내용은 클러스터를 생성할 때 구성 제공 단원을 참조하십시오.
- 보안 구성의 로컬 디스크 암호화 및 S3 암호화 개선(2019년 8월 5일)
  - 보안 구성 설정에서 로컬 디스크 암호화 설정과 Amazon S3 암호화 설정이 분리되었습니다.
  - EBS 암호화를 활성화하는 옵션이 5.24.0 이상 릴리스에 추가되었습니다. 이 옵션을 선택하면 스토리지 볼륨뿐만 아니라 루트 디바이스 볼륨도 암호화됩니다. 이전 버전에서는 사용자 지정 AMI를 사용하여 루트 디바이스 볼륨을 암호화해야 했습니다.
  - 자세한 내용은 Amazon EMR 관리 안내서에서 암호화 옵션을 참조하세요.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:
  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.

- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.24.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.24.1에 대한 정보가 포함됩니다. 변경 사항은 5.24.0에 관련됩니다.

최초 릴리스 날짜: 2019년 6월 26일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.24.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.24.0에 대한 정보가 포함됩니다. 변경 사항은 5.23.0에 관련됩니다.

최초 릴리스 날짜: 2019년 6월 11일

최종 업데이트 날짜: 2019년 8월 5일

업그레이드

- Flink 1.8.0
- Hue 4.4.0
- JupyterHub 0.9.6
- Livy 0.6.0
- MxNet 1.4.0
- Presto 0.219
- Spark 2.4.2
- AWS SDK for Java 1.11.546
- 커넥터 및 드라이버:

- DynamoDB 커넥터 4.9.0
- MariaDB 커넥터 2.4.1
- Amazon Redshift JDBC 드라이버 1.2.27.1051

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - 파티션을 동적으로 삭제하는 최적화가 추가되었습니다. 최적화는 기본적으로 비활성화되어 있습니다. 활성화하려면 Spark 구성 파라미터 `spark.sql.dynamicPartitionPruning.enabled`를 `true`로 설정하십시오.
  - `INTERSECT` 쿼리의 성능을 개선했습니다. 이 최적화는 기본적으로 비활성화되어 있습니다. 활성화하려면 Spark 구성 파라미터 `spark.sql.optimizer.distinctBeforeIntersect.enabled`를 `true`로 설정하십시오.
  - 동일한 관계를 사용하는 집계로 스칼라 하위 쿼리를 평면화하는 최적화가 추가되었습니다. 최적화는 기본적으로 비활성화되어 있습니다. 활성화하려면 Spark 구성 파라미터 `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled`를 `true`로 설정하십시오.
  - 전체 스테이지 코드 생성이 개선되었습니다.

  자세한 내용은 Spark 성능 최적화를 참조하십시오.
- 보안 구성의 로컬 디스크 암호화 및 S3 암호화 개선(2019년 8월 5일)
  - 보안 구성 설정에서 로컬 디스크 암호화 설정과 Amazon S3 암호화 설정이 분리되었습니다.
  - EBS 암호화를 활성화하는 옵션이 추가되었습니다. 이 옵션을 선택하면 스토리지 볼륨뿐만 아니라 루트 디바이스 볼륨도 암호화됩니다. 이전 버전에서는 사용자 지정 AMI를 사용하여 루트 디바이스 볼륨을 암호화해야 했습니다.
  - 자세한 내용은 Amazon EMR 관리 안내서에서 암호화 옵션을 참조하세요.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일

다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.23.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.23.0에 대한 정보가 포함됩니다. 변경 사항은 5.22.0에 관련됩니다.

최초 릴리스 날짜: 2019년 4월 1일

최종 업데이트 날짜: 2019년 4월 30일

업그레이드

- AWS SDK for Java 1.11.519

새로운 특성

- (2019년 4월 30일) Amazon EMR 5.23.0 이상에서 3개의 프라이머리 노드를 포함하는 클러스터를 시작하고 YARN Resource Manager, HDFS NameNode, Spark, Hive, Ganglia와 같은 애플리케이션의 고가용성을 지원할 수 있습니다. 이 기능을 사용하면 프라이머리 노드가 더 이상 잠재적 단일 장애 지점이 아닙니다. 프라이머리 노드 중 하나에 장애가 발생할 경우, Amazon EMR이 자동으로 대

기 프라이머리 노드로 장애 조치하고 장애가 발생한 프라이머리 노드를 동일한 구성 및 부트스트랩 작업을 갖는 새로운 프라이머리 노드로 교체합니다. 자세한 내용은 [프라이머리 노드 계획 및 구성](#)을 참조하세요.

알려진 문제

- Tez UI(Amazon EMR 릴리스 5.26.0에서 수정됨)

  Tez UI는 여러 프라이머리 노드를 포함하는 EMR 클러스터에서 작동하지 않습니다.
- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.
    2. 텍스트 편집기를 사용하여 `hue.ini` 파일을 엽니다. 예시:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. `appblacklist`라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

       ```
       appblacklist = search
       ```
    4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

       ```
       sudo stop hue; sudo start hue
       ```
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일

다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
>
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.22.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.22.0에 대한 정보가 포함됩니다. 변경 사항은 5.21.0에 관련됩니다.

> **Important**
>
> Amazon EMR 릴리스 5.22.0부터 Amazon EMR은 AWS 서명 버전 4만 사용하여 Amazon S3에 대한 요청을 인증합니다. 이전 Amazon EMR 릴리스는 경우에 따라 AWS 서명 버전 4만 사용된다는 릴리스 정보가 표시되지 않는 한 서명 버전 2를 사용합니다. 자세한 내용은 Amazon Simple Storage Service 개발자 안내서의 [요청 인증(AWS 서명 버전 4)](#) 및 [요청 인증(AWS 서명 버전 2)](#)을 참조하세요.

최초 릴리스 날짜: 2019년 3월 20일

업그레이드

- Flink 1.7.1
- HBase 1.4.9
- Oozie 5.1.0
- Phoenix 4.14.1
- Zeppelin 0.8.1
- 커넥터 및 드라이버:
  - DynamoDB 커넥터 4.8.0
  - MariaDB Connector 2.2.6
  - Amazon Redshift JDBC 드라이버 1.2.20.1043

새로운 특성

- EBS 전용 스토리지를 사용하는 EC2 인스턴스 유형에 대한 기본 EBS 구성을 수정했습니다. Amazon EMR 릴리스 5.22.0 이상을 사용하여 클러스터를 생성할 때 EBS 스토리지의 기본적인 크기가 인스턴스 크기에 따라 증가합니다. 뿐만 아니라 증가한 스토리지를 여러 볼륨에 분할하여 향상된 IOPS 성능을 제공합니다. 다른 EBS 인스턴스 스토리지 구성을 사용하려는 경우 EMR 클러스터를 생성하거나 기존 클러스터에 노드를 추가할 때 이 구성을 지정할 수 있습니다. 각 인스턴스 유형에 기본적으로 할당된 스토리지 크기와 볼륨 수에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 인스턴스의 기본 EBS 스토리지를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - YARN의 Spark에 새로운 구성 속성 `spark.yarn.executor.memoryOverheadFactor`를 추가했습니다. 이 속성의 값은 최소 384MB의 메모리 오버헤드 값을 실행기 메모리의 백분율로 설정하는 확장 팩터입니다. 메모리 오버헤드가 `spark.yarn.executor.memoryOverhead`를 사용하여 명시적으로 설정된 경우 이 속성은 아무 효과가 없습니다. 기본값은 18.75%를 나타내는 `0.1875`입니다. Amazon EMR의 기본값은 실행기 메모리 오버헤드를 위해 Spark에서 내부적으로 설정한 기본값 10%보다 더 많은 공간을 YARN 컨테이너에 남깁니다. TPC-DS 벤치마크 결과 Amazon EMR 기본값, 18.75%로 설정하면 메모리 관련 오류가 경험적으로 적었습니다.
  - 성능 개선을 위해 SPARK-26316을 백포트했습니다.
- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수

있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.

알려진 문제

- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.
    2. 텍스트 편집기를 사용하여 `hue.ini` 파일을 엽니다. 예시:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

    3. `appblacklist`라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

       ```
       appblacklist = search
       ```

    4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

       ```
       sudo stop hue; sudo start hue
       ```

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.21.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.21.1에 대한 정보가 포함됩니다. 변경 사항은 5.21.0에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 18일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

알려진 문제

- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

해결 방법:

- 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
- 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.21.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.21.0에 대한 정보가 포함됩니다. 변경 사항은 5.20.0에 관련됩니다.

최초 릴리스 날짜: 2019년 2월 18일

최종 업데이트 날짜: 2019년 4월 3일

업그레이드

- Flink 1.7.0
- Presto 0.215
- AWS SDK for Java 1.11.479

새로운 특성

- (2019년 4월 3일) Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘

솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- Zeppelin
  - ZEPPELIN-3878을 백포트했습니다.

알려진 문제

- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.
    2. 텍스트 편집기를 사용하여 `hue.ini` 파일을 엽니다. 예시:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```
    3. `appblacklist`라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

       ```
       appblacklist = search
       ```
    4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

       ```
       sudo stop hue; sudo start hue
       ```
- Tez의 Hive 실행 시간 비교
  - 이 문제는 Amazon EMR 5.22.0에서 해결되었습니다.

클러스터 프라이머리 노드에 대한 SSH 연결을 통해 http://*MasterDNS*:8080/tez-ui의 Tez UI에 연결할 때 'Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled(어댑터 작동이 실패했습니다. - 타임라인 서버(ATS)가 범위를 벗어났습니다. 다운되었거나 CORS가 활성화되어 있지 않습니다.)"라는 오류가 나타나거나, 작업 중 예기치 않게 N/A(해당 사항 없음)가 표시됩니다.

이 오류는 Tez UI가 프라이머리 노드의 호스트 이름 대신 `localhost`를 사용하여 YARN 타임라인 서버에 요청할 때 발생합니다. 이 문제를 해결하기 위해 스크립트를 부트스트랩 작업 또는 단계로 실행할 수 있습니다. 스크립트는 Tez `configs.env` 파일의 호스트 이름을 업데이트합니다. 스크립트에 대한 자세한 내용과 위치를 보려면 부트스트랩 지침을 참조하십시오.

- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수 있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

  ```
  kinit -kt <keytab_file> <principal>
  ```

  일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.20.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.20.0에 대한 정보가 포함됩니다. 변경 사항은 5.19.0에 관련됩니다.

최초 릴리스 날짜: 2018년 12월 18일

마지막 업데이트 날짜: 2019년 1월 22일

업그레이드

- Flink 1.6.2
- HBase 1.4.8
- Hive 2.3.4
- Hue 4.3.0
- MXNet 1.3.1
- Presto 0.214
- Spark 2.4.0
- TensorFlow 1.12.0
- Tez 0.9.1
- AWS SDK for Java 1.11.461

새로운 특성

- (2019년 1월 22일) Amazon EMR의 Kerberos가 외부 KDC의 인증 보안 주체를 지원하도록 개선되었습니다. 이는 여러 클러스터가 하나의 외부 KDC를 공유할 수 있기 때문에 보안 주체 관리를 중앙 집중화합니다. 또한 외부 KDC는 Active Directory 도메인과 교차 영역 신뢰를 가질 수 있습니다. 이를 통해 모든 클러스터는 Active Directory의 보안 주체를 인증할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Kerberos 인증 사용을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- Amazon EMR의 기본 Amazon Linux AMI
  - Python3 패키지가 python 3.4에서 3.6으로 업그레이드되었습니다.
- EMRFS S3 최적화 커미터
  - EMRFS S3 최적화 커미터가 이제 기본적으로 활성화되며 쓰기 성능을 향상시킵니다. 자세한 내용은 EMRFS S3 최적화 커미터 사용 단원을 참조하십시오.
- Hive
  - HIVE-16686를 백포트했습니다.
- Glue와 Spark 및 Hive
  - EMR 5.20.0 이상에서는 AWS Glue 데이터 카탈로그를 메타스토어로 사용할 때 Spark 및 Hive에 대해 병렬 파티션 정리가 자동으로 활성화됩니다. 이 변경 사항은 여러 요청을 병렬로 실행하여 파티션을 검색함으로써 쿼리 계획 시간을 크게 단축합니다. 동시에 실행할 수 있는 총 세그먼트 수는 1~10입니다. 기본값은 5이며, 또한 권장 설정입니다. `hive-site` 구성 분류에서 `aws.glue.partition.num.segments` 속성을 지정하여 이 설정을 변경할 수 있습니다. 조절이 발생할 경우 값을 1로 변경하여 기능을 끌 수 있습니다. 자세한 내용은 AWS Glue 그먼트 구조를 참조하세요.

알려진 문제

- Hue(Amazon EMR 릴리스 5.24.0에서 수정됨)
  - Amazon EMR에서 실행되는 Hue는 Solr을 지원하지 않습니다. Amazon EMR 릴리스 5.20.0부터 잘못된 구성 문제로 인해 Solr이 활성화되고 다음과 유사한 무해한 오류 메시지가 표시됩니다.

    ```
    Solr server could not be contacted properly:
    HTTPConnectionPool('host=ip-xx-xx-xx-xx.ec2.internal',
    port=1978): Max retries exceeded with url: /solr/admin/info/
    system?user.name=hue&doAs=administrator&wt=json (Caused by
    NewConnectionError(': Failed to establish a new connection: [Errno 111]
    Connection refused',))
    ```

    Solr 오류 메시지를 표시하지 않으려면

    1. SSH를 사용하여 프라이머리 노드 명령줄에 연결합니다.
    2. 텍스트 편집기를 사용하여 `hue.ini` 파일을 엽니다. 예시:

       ```
       sudo vim /etc/hue/conf/hue.ini
       ```

3. appblacklist라는 단어를 검색하고 해당 줄을 다음과 같이 수정합니다.

```
appblacklist = search
```

4. 변경 내용을 저장하고 다음 예와 같이 Hue를 다시 시작합니다.

```
sudo stop hue; sudo start hue
```

- Tez의 Hive 실행 시간 비교
  - 이 문제는 Amazon EMR 5.22.0에서 해결되었습니다.

    클러스터 프라이머리 노드에 대한 SSH 연결을 통해 http://*MasterDNS*:8080/tez-ui의 Tez UI에 연결할 때 'Adapter operation failed - Timeline server (ATS) is out of reach. Either it is down, or CORS is not enabled(어댑터 작동이 실패했습니다. - 타임라인 서버(ATS)가 범위를 벗어났습니다. 다운되었거나 CORS가 활성화되어 있지 않습니다.)"라는 오류가 나타나거나, 작업 중 예기치 않게 N/A(해당 사항 없음)가 표시됩니다.

    이 오류는 Tez UI가 프라이머리 노드의 호스트 이름 대신 `localhost`를 사용하여 YARN 타임라인 서버에 요청할 때 발생합니다. 이 문제를 해결하기 위해 스크립트를 부트스트랩 작업 또는 단계로 실행할 수 있습니다. 스크립트는 Tez `configs.env` 파일의 호스트 이름을 업데이트합니다. 스크립트에 대한 자세한 내용과 위치를 보려면 부트스트랩 지침을 참조하십시오.
- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수 있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.
- 여러 프라이머리 노드와 Kerberos 인증을 사용하는 클러스터에서 알려진 문제

  Amazon EMR 릴리스 5.20.0 이상에서 여러 프라이머리 노드와 Kerberos 인증을 사용하여 클러스터를 실행하는 경우, 클러스터를 일정 시간 실행한 후 스케일 다운 또는 단계 제출과 같은 클러스터 작업에 문제가 발생할 수 있습니다. 기간은 정의한 Kerberos 티켓 유효 기간에 따라 다릅니다. 스케일 다운 문제는 제출한 자동 스케일 다운 및 명시적 스케일 다운 요청 모두에 영향을 미칩니다. 이외의 다른 클러스터 작업도 영향을 받을 수 있습니다.

  해결 방법:

  - 여러 프라이머리 노드가 있는 EMR 클러스터의 리드 프라이머리 노드에 대한 hadoop 사용자로 SSH 연결합니다.
  - 다음 명령을 실행하여 hadoop 사용자의 Kerberos 티켓을 갱신합니다.

```
kinit -kt <keytab_file> <principal>
```

일반적으로 keytab 파일은 `/etc/hadoop.keytab`에 있으며, 보안 주체는 `hadoop/<hostname>@<REALM>` 양식입니다.

> **Note**
> 이 해결 방법은 Kerberos 티켓이 유효한 기간에 작동합니다. 이 기간은 기본적으로 10시간이지만 Kerberos 설정으로 구성할 수 있습니다. Kerberos 티켓이 만료되면 위 명령을 다시 실행해야 합니다.

릴리스 5.19.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.19.0에 대한 정보가 포함됩니다. 변경 사항은 5.18.0에 관련됩니다.

최초 릴리스 날짜: 2018년 11월 7일

최종 업데이트: 2018년 11월 19일

업그레이드

- 하둡 2.8.5
- Flink 1.6.1
- JupyterHub 0.9.4
- MXNet 1.3.0
- Presto 0.212
- TensorFlow 1.11.0
- Zookeeper 3.4.13
- AWS SDK for Java 1.11.433

새로운 특성

- (2018년 11월 19일) EMR Notebooks는 Jupyter Notebook 기반의 관리형 환경입니다. PySpark, Spark SQL, Spark R 및 Scala에 대한 Spark 매직 커널을 지원합니다. EMR Notebooks는 Amazon

EMR 릴리스 5.18.0 이상을 사용하여 생성한 클러스터와 함께 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR Notebooks 사용을 참조하세요.

- Spark 및 EMRFS를 사용하여 Parquet 파일을 작성할 때 EMRFS S3 최적화 커미터를 사용할 수 있습니다. 이 커미터는 쓰기 성능을 향상시킵니다. 자세한 내용은 EMRFS S3 최적화 커미터 사용 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- YARN
  - 코어 노드에서 실행할 때 애플리케이션 마스터 프로세스를 제한하는 로직이 수정되었습니다. 이 기능은 이제 YARN 노드 레이블 기능과 `yarn-site`의 속성 및 `capacity-scheduler` 구성 분류를 사용합니다. 자세한 정보는 https://docs.aws.amazon.com/emr/latest/ManagementGuide/emr-plan-instances-guidelines.html#emr-plan-spot-YARN. 섹션을 참조하세요.
- Amazon EMR의 기본 Amazon Linux AMI
  - `ruby18`, `php56`, `gcc48`이 이제 기본적으로 설치되지 않습니다. 필요할 경우 `yum`을 사용하여 설치할 수 있습니다.
  - aws-java-sdk ruby gem이 이제 더 이상 기본적으로 설치되지 않습니다. 필요할 경우 `gem install aws-sdk`를 사용하여 설치할 수 있습니다. 특정 구성 요소도 설치할 수 있습니다. 예를 들어 `gem install aws-sdk-s3`입니다.

알려진 문제

- EMR Notebooks - 일부 상황에서 노트북 편집기가 여러 개 열려 있는 경우 노트북 편집기가 클러스터에 연결할 수 없는 것처럼 보일 수 있습니다. 이 경우 브라우저 쿠키를 지우고 노트북 편집기를 다시 여십시오.
- CloudWatch ContainerPending 지표 및 자동 조정 - (5.20.0에서 수정됨) Amazon EMR은 `ContainerPending`에 대해 음수 값을 생성할 수 있습니다. 자동 조정 규칙에 `ContainerPending`을 사용할 경우 자동 조정이 예상대로 작동하지 않습니다. 자동 조정에 `ContainerPending`을 사용하지 마십시오.
- Amazon EMR 버전 5.19.0, 5.20.0 및 5.21.0에서 YARN 노드 레이블은 HDFS 디렉터리에 저장됩니다. 이로 인해 경우에 따라 코어 노드 시작이 지연되어 클러스터가 시간 초과되고 시작이 실패할 수 있습니다. Amazon EMR 5.22.0부터 이 문제가 해결되었습니다. YARN 노드 레이블은 각 클러스터 노드의 로컬 디스크에 저장되어 HDFS에 종속되지 않습니다.

릴리스 5.18.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.18.0에 대한 정보가 포함됩니다. 변경 사항은 5.17.0에 관련됩니다.

최초 릴리스 날짜: 2018년 10월 24일

업그레이드

- Flink 1.6.0
- HBase 1.4.7
- Presto 0.210
- Spark 2.3.2
- Zeppelin 0.8.0

새로운 특성

- Amazon EMR 5.18.0부터 Amazon EMR 아티팩트 리포지토리를 사용하여 특정 Amazon EMR 릴리스 버전에서 사용할 수 있는 정확한 버전의 라이브러리와 종속 항목에 대해 작업 코드를 빌드할 수 있습니다. 자세한 내용은 Amazon EMR 아티팩트 리포지토리를 사용하여 종속성 확인 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- Hive
  - S3 Select에 대한 지원을 추가했습니다. 자세한 내용은 S3 Select와 함께 Hive를 사용하여 성능 향상 단원을 참조하십시오.
- Presto
  - S3 Select Pushdown에 대한 지원을 추가했습니다. 자세한 내용은 S3 Select Pushdown과 함께 Presto를 사용하여 성능 향상 단원을 참조하십시오.
- Spark
  - Spark에 대한 기본 log4j 구성이 Spark 스트리밍 작업에 대한 컨테이너 로그를 1시간마다 넘기도록 변경되었습니다. 따라서 오래 실행되는 Spark 스트리밍 작업에 대한 로그가 삭제되지 않습니다.

릴리스 5.17.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.17.1에 대한 정보가 포함됩니다. 변경 사항은 5.17.0에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 18일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

릴리스 5.17.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.17.0에 대한 정보가 포함됩니다. 변경 사항은 5.16.0에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 30일

업그레이드

- Flink 1.5.2
- HBase 1.4.6
- Presto 0.206

새로운 특성

- Tensorflow에 대한 지원을 추가했습니다. 자세한 내용은 TensorFlow 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- JupyterHub
  - Amazon S3의 노트북 지속성에 대한 지원을 추가했습니다. 자세한 내용은 Amazon S3에서 노트북의 지속성 구성 단원을 참조하십시오.
- Spark
  - S3 Select에 대한 지원을 추가했습니다. 자세한 내용은 S3 Select와 함께 Spark를 사용하여 쿼리 성능 향상 단원을 참조하십시오.

- Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0의 CloudWatch 지표 및 자동 조정 기능 관련 문제를 해결했습니다.

알려진 문제

- Kerberos 인증을 사용하는 클러스터를 생성하고 Livy를 설치하면 단순 인증이 활성화되지 않았다는 오류가 발생하면서 Livy가 실패합니다. Livy 서버를 재부팅하면 문제가 해결됩니다. 해결 방법으로, 클러스터 생성 중에 프라이머리 노드에서 `sudo restart livy-server`를 실행하는 단계를 추가합니다.
- 생성 날짜가 2018-08-11인 Amazon Linux AMI를 기반으로 사용자 지정 Amazon Linux AMI를 사용하면 Oozie 서버가 시작에 실패합니다. Oozie를 사용하는 경우 생성 날짜가 다른 Amazon Linux AMI ID를 기반으로 사용자 지정 AMI를 생성하십시오. 다음 AWS CLI 명령을 사용하여 릴리스 날짜와 함께 2018.03 버전의 모든 HVM Amazon Linux AMIs에 대한 이미지 IDs 목록을 반환할 수 있으므로 적절한 Amazon Linux AMI를 기본으로 선택할 수 있습니다. MyRegion을 사용자의 리전 식별자(예: us-west-2)로 바꿉니다.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

릴리스 5.16.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.16.0에 대한 정보가 포함됩니다. 변경 사항은 5.15.0에 관련됩니다.

최초 릴리스 날짜: 2018년 7월 19일

업그레이드

- Hadoop 2.8.4
- Flink 1.5.0
- Livy 0.5.0
- MXNet 1.2.0
- Phoenix 4.14.0
- Presto 0.203
- Spark 2.3.1

- AWS SDK for Java 1.11.336
- CUDA 9.2
- Redshift JDBC 드라이버 1.2.15.1025

변경 사항, 향상된 기능 및 해결된 문제

- HBase
  - HBASE-20723 백포트
- Presto
  - LDAP 인증을 지원하도록 구성을 변경합니다. 자세한 내용은 Amazon EMR 기반 Presto에 대한 LDAP 인증 사용 단원을 참조하십시오.
- Spark
  - Amazon EMR 릴리스 5.16.0부터 사용할 수 있는 Apache Spark 버전 2.3.1은 CVE-2018-8024 및 CVE-2018-1334를 처리합니다. Spark의 이전 버전을 Spark 버전 2.3.1 이상으로 마이그레이션하는 것이 좋습니다.

알려진 문제

- 이 릴리스 버전은 c1.medium 또는 m1.small 인스턴스 유형을 지원하지 않습니다. 둘 중 어느 쪽이든 이런 인스턴스 유형을 사용하는 클러스터는 시작되지 않습니다. 차선책으로 다른 인스턴스 유형을 지정하거나 다른 릴리스 버전을 사용하십시오.
- Kerberos 인증을 사용하는 클러스터를 생성하고 Livy를 설치하면 단순 인증이 활성화되지 않았다는 오류가 발생하면서 Livy가 실패합니다. Livy 서버를 재부팅하면 문제가 해결됩니다. 해결 방법으로, 클러스터 생성 중에 프라이머리 노드에서 `sudo restart livy-server`를 실행하는 단계를 추가합니다.
- 프라이머리 노드가 재부팅되거나 인스턴스 컨트롤러가 다시 시작된 후에는 CloudWatch 지표가 수집되지 않으며 Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0에서는 자동 조정 기능을 사용할 수 없습니다. 이 문제는 Amazon EMR 5.17.0에서 수정되었습니다.

릴리스 5.15.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.15.0에 대한 정보가 포함됩니다. 변경 사항은 5.14.0에 관련됩니다.

최초 릴리스 날짜: 2018년 6월 21일

업그레이드

- HBase를 1.4.4로 업그레이드
- Hive를 2.3.3으로 업그레이드
- Hue를 4.2.0으로 업그레이드
- Oozie를 5.0.0로 업그레이드
- Zookeeper를 3.4.12로 업그레이드
- AWS SDK를 1.11.333으로 업그레이드

변경 사항, 향상된 기능 및 해결된 문제

- Hive
  - HIVE-18069를 백포트했습니다.
- Hue
  - Kerberos 사용 시 Livy로 올바로 인증하도록 Hue를 업데이트했습니다. 이제 Livy가 Amazon EMR에서 Kerberos를 사용할 때 지원됩니다.
- JupyterHub
  - Amazon EMR이 LDAP 클라이언트 라이브러리를 기본으로 설치하도록 JupyterHub를 업데이트했습니다.
  - 자체 서명 인증서를 생성하는 스크립트 오류를 수정했습니다.

알려진 문제

- 이 릴리스 버전은 c1.medium 또는 m1.small 인스턴스 유형을 지원하지 않습니다. 둘 중 어느 쪽이든 이런 인스턴스 유형을 사용하는 클러스터는 시작되지 않습니다. 차선책으로 다른 인스턴스 유형을 지정하거나 다른 릴리스 버전을 사용하십시오.
- 프라이머리 노드가 재부팅되거나 인스턴스 컨트롤러가 다시 시작된 후에는 CloudWatch 지표가 수집되지 않으며 Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0에서는 자동 조정 기능을 사용할 수 없습니다. 이 문제는 Amazon EMR 5.17.0에서 수정되었습니다.

릴리스 5.14.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.14.1에 대한 정보가 포함됩니다. 변경 사항은 5.14.0에 관련됩니다.

최초 릴리스 날짜: 2018년 10월 17일

잠재적 보안 취약성을 해결하기 위해 Amazon EMR의 기본 AMI를 업데이트했습니다.

릴리스 5.14.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.14.0에 대한 정보가 포함됩니다. 변경 사항은 5.13.0에 관련됩니다.

최초 릴리스 날짜: 2018년 6월 4일

업그레이드

- Apache Flink를 1.4.2로 업그레이드
- Apache MXnet을 1.1.0으로 업그레이드
- Apache Sqoop를 1.4.7로 업그레이드

새로운 특성

- JupyterHub 지원을 추가했습니다. 자세한 내용은 JupyterHub 단원을 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- EMRFS
  - Amazon S3에 대한 요청의 userAgent 문자열이 간접 호출하는 보안 주체의 사용자 및 그룹 정보를 포함하도록 업데이트되었습니다. 이는 보다 포괄적인 요청 추적을 위해 AWS CloudTrail 로그와 함께 사용할 수 있습니다.
- HBase
  - 포함된 HBASE-20447은 특히 분할된 리전과 함께 캐시 문제를 야기할 수 있는 문제를 해결합니다.
- MXnet
  - OpenCV 라이브러리를 추가했습니다.
- Spark
  - Spark가 EMRFS를 사용하여 Amazon S3 위치에 Parquet 파일을 쓸 때, FileOutputCommitter 알고리즘이 버전 1 대신 버전 2를 사용하도록 업데이트되었습니다. 따라서 이름 변경 횟수가 감소하여 애플리케이션 성능이 개선됩니다. 다음 애플리케이션에는 이 변경이 영향을 미치지 않습니다.
    - Spark 이외의 애플리케이션

- HDFS(여전히 버전 1의 FileOutputCommitter를 사용)와 같이 다른 파일 시스템에 쓰는 애플리케이션
- 텍스트 또는 csv 같이 이미 EMRFS 직접 쓰기를 사용하는 다른 출력 형식을 사용하는 애플리케이션

알려진 문제

- JupyterHub
  - 클러스터를 생성할 때 구성 분류를 사용하여 JupyterHub 및 개별 Jupyter Notebook을 설정하는 것은 지원되지 않습니다. 각 사용자에 대해 수동으로 jupyterhub_config.py 파일 및 jupyter_notebook_config.py 파일을 편집하십시오. 자세한 내용은 [JupyterHub 구성](#) 단원을 참조하십시오.
  - JupyterHub는 프라이빗 서브넷 내의 클러스터에서 시작되지 않으며 `Error: ENOENT: no such file or directory, open '/etc/jupyter/conf/server.crt'` 메시지와 함께 실패합니다. 이는 자체 서명된 인증서를 생성하는 스크립트의 오류에 의해 야기됩니다. 다음 차선책을 사용하여 자체 서명된 인증서를 생성합니다. 모든 명령은 프라이머리 노드에 연결되어 있는 동안 실행됩니다.

    1. 컨테이너에서 프라이머리 노드로 인증서 생성 스크립트를 복사합니다.

    ```
    sudo docker cp jupyterhub:/tmp/gen_self_signed_cert.sh ./
    ```

    2. 아래와 같이 텍스트 편집기를 사용하여 23행을 변경하고 퍼블릭 호스트 이름을 로컬 호스트 이름으로 변경합니다.

    ```
    local hostname=$(curl -s $EC2_METADATA_SERVICE_URI/local-hostname)
    ```

    3. 스크립트를 실행하여 자체 서명된 인증서를 생성합니다.

    ```
    sudo bash ./gen_self_signed_cert.sh
    ```

    4. 스크립트가 생성하는 인증서 파일을 `/etc/jupyter/conf/` 디렉터리로 이동합니다.

    ```
    sudo mv /tmp/server.crt /tmp/server.key /etc/jupyter/conf/
    ```

    JupyterHub의 재시작과 200 응답 코드의 반환을 확인하기 위해 `jupyter.log` 파일을 `tail`할 수 있습니다. 예시:

```
tail -f /var/log/jupyter/jupyter.log
```

그러면 다음과 비슷한 응답이 돌아옵니다.

```
# [I 2018-06-14 18:56:51.356 JupyterHub app:1581] JupyterHub is now running at
 https://:9443/
# 19:01:51.359 - info: [ConfigProxy] 200 GET /api/routes
```

- 프라이머리 노드가 재부팅되거나 인스턴스 컨트롤러가 다시 시작된 후에는 CloudWatch 지표가 수집되지 않으며 Amazon EMR 버전 5.14.0, 5.15.0 또는 5.16.0에서는 자동 조정 기능을 사용할 수 없습니다. 이 문제는 Amazon EMR 5.17.0에서 수정되었습니다.

릴리스 5.13.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.13.0에 대한 정보가 포함됩니다. 변경 사항은 5.12.0에 관련됩니다.

업그레이드

- Spark를 2.3.0으로 업그레이드
- HBase를 1.4.2로 업그레이드
- Presto를 0.194로 업그레이드
- 1.11.297 AWS SDK for Java 로 업그레이드됨

변경 사항, 향상된 기능 및 해결된 문제

- Hive
  - HIVE-15436을 백포트했습니다. Hive API를 뷰만 반환하도록 개선했습니다.

알려진 문제

- 현재 MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.

릴리스 5.12.2

다음 릴리스 정보에는 Amazon EMR 릴리스 5.12.2에 대한 정보가 포함됩니다. 변경 사항은 5.12.1에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

릴리스 5.12.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.12.1에 대한 정보가 포함됩니다. 변경 사항은 5.12.0에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

릴리스 5.12.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.12.0에 대한 정보가 포함됩니다. 변경 사항은 5.11.1에 관련됩니다.

업그레이드

- AWS SDK for Java 1.11.238 " 1.11.267. 자세한 내용은 GitHub에서 AWS SDK for Java Change Log를 참조하세요.
- Hadoop 2.7.3 ⇒ 2.8.3. 자세한 내용은 Apache Hadoop Releases를 참조하십시오.
- Flink 1.3.2 ⇒ 1.4.0. 자세한 내용은 Apache Flink 1.4.0 Release Announcement를 참조하십시오.
- HBase 1.3.1 ⇒ 1.4.0. 자세한 내용은 HBase Release Announcement를 참조하십시오.
- Hue 4.0.1 ⇒ 4.1.0. 자세한 내용은 릴리스 정보를 참조하십시오.
- MxNet 0.12.0 ⇒ 1.0.0. 자세한 내용은 GitHub에서 MXNet Change Log를 참조하십시오.

- Presto 0.187 ⇒ 0.188. 자세한 내용은 릴리스 정보를 참조하십시오.

변경 사항, 향상된 기능 및 해결된 문제

- Hadoop
  - `yarn.resourcemanager.decommissioning.timeout` 속성을 `yarn.resourcemanager.nodemanager-graceful-decommission-timeout-secs`로 변경했습니다. 이 속성을 사용하여 클러스터 축소를 사용자 지정할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 스케일 다운을 참조하세요.
  - Hadoop CLI는 직접 복사를 지정하는 -d(복사) 명령에 `cp` 옵션을 추가했습니다. 이를 사용하면 중간 `.COPYING` 파일을 생성하지 않아도 되므로 Amazon S3 사이에서 데이터를 더 빠르게 복사할 수 있습니다. 자세한 내용은 HADOOP-12384를 참조하십시오.
- Pig
  - Pig 환경 속성의 구성을 단순화하는 `pig-env` 구성 분류를 추가했습니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.
- Presto
  - Presto `presto-connector-redshift` 구성 파일에서 값을 구성하는 데 사용할 수 있는 `redshift.properties` 구성 분류를 추가했습니다. 자세한 내용은 Presto 문서의 Redshift Connector 및 애플리케이션 구성 단원을 참조하십시오.
  - EMRFS에 대한 Presto 지원이 추가되었으며 이것이 기본 구성입니다. 이전 Amazon EMR 릴리스에서는 유일한 옵션인 PrestoS3FileSystem을 사용했습니다. 자세한 내용은 EMRFS 및 PrestoS3FileSystem 구성 단원을 참조하십시오.

  > **Note**
  >
  > Amazon EMR 버전 5.12.0을 사용하여 Amazon S3에서 기본 데이터를 쿼리하는 경우 Presto 오류가 발생할 수 있습니다. 이는 Presto가 `emrfs-site.xml`에서 구성 분류 값을 가져오지 못하기 때문입니다. 이 문제를 해결하려면 `usr/lib/presto/plugin/hive-hadoop2/` 아래에 `emrfs` 하위 디렉터리를 생성하고 `usr/lib/presto/plugin/hive-hadoop2/emrfs`에서 기존 `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` 파일에 대한 기호 링크를 생성합니다. 그런 다음, presto-server 프로세스를 다시 시작합니다(`sudo presto-server stop` 및 `sudo presto-server start`의 순서대로)

- Spark

- SPARK-22036: BigDecimal multiplication sometimes returns null을 백포트했습니다.

알려진 문제

- MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.
- 클러스터 노드에 R이 기본적으로 설치되어 있지 않으므로 사용자 지정 AMI를 사용하여 만든 클러스터에는 SparkR을 사용할 수 없습니다.

릴리스 5.11.3

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.3에 대한 정보가 포함됩니다. 변경 사항은 5.11.2에 관련됩니다.

최초 릴리스 날짜: 2019년 7월 18일

변경 사항, 향상된 기능 및 해결된 문제

- TCP SACK 서비스 거부 문제(AWS-2019-005)를 비롯한 중요한 Linux 커널 보안 업데이트를 포함하도록 Amazon EMR의 기본 Amazon Linux AMI를 업데이트했습니다.

릴리스 5.11.2

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.2에 대한 정보가 포함됩니다. 변경 사항은 5.11.1에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

릴리스 5.11.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.11.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

알려진 문제

- MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.
- Hive 2.3.2에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 Statistics in Hive를 참조하십시오.

릴리스 5.11.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.11.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.10.0 릴리스에 관련됩니다.

업그레이드

이 릴리스에서는 다음 버전을 포함하도록 다음 애플리케이션과 구성 요소를 업그레이드했습니다.

- Hive 2.3.2
- Spark 2.2.1
- SDK for Java 1.11.238

새로운 특성

- Spark

- 스팟 인스턴스 사용 시 Spark 폐기를 개선하는 `spark.decommissioning.timeout.threshold` 설정이 추가되었습니다. 자세한 내용은 노드 서비스 해제 동작 구성 단원을 참조하십시오.
- Amazon SageMaker Spark 및 Amazon SageMaker와 Spark의 통합을 위한 관련 종속 항목을 설치하는 `aws-sagemaker-spark-sdk` 구성 요소가 Spark에 추가되었습니다. Amazon SageMaker Spark를 사용하면 Amazon SageMaker 단계를 사용하여 Spark 기계 학습(ML) 파이프라인을 생성할 수 있습니다. 자세한 내용은 GitHub의 SageMaker Spark Readme와 Amazon SageMaker 개발자 안내서의 Using Apache Spark with Amazon SageMaker를 참조하세요.

알려진 문제

- MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.
- Hive 2.3.2에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 Statistics in Hive를 참조하십시오.

릴리스 5.10.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.10.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.9.0 릴리스에 관련됩니다.

업그레이드

이 릴리스에서는 다음 버전을 포함하도록 다음 애플리케이션과 구성 요소를 업그레이드했습니다.

- AWS SDK for Java 1.11.221
- Hive 2.3.1
- Presto 0.187

새로운 특성

- Kerberos 인증에 대한 지원이 추가되었습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Kerberos 인증 사용을 참조하세요.
- Amazon S3에 대한 EMRFS 요청을 위해 IAM 역할에 대한 지원을 추가했습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon S3에 대한 EMRFS 요청을 위해 IAM 역할 구성을 참조하세요.
- GPU 기반의 P2 및 P3 인스턴스 유형에 대한 지원을 추가했습니다. 자세한 내용은 Amazon EC2 P2 인스턴스 및 Amazon EC2 P3 인스턴스를 참조하세요. 기본적으로 이러한 유형의 인스턴스에 NVIDIA 드라이버 384.81 및 CUDA 드라이버 9.0.176가 설치됩니다.
- Apache MXNet 지원이 추가되었습니다.

변경 사항, 향상된 기능 및 해결된 문제

- Presto
  - Glue 데이터 카탈로그를 기본 Hive AWS 메타스토어로 사용하기 위한 지원이 추가되었습니다. 자세한 내용은 AWS Glue 데이터 카탈로그에서 Presto 사용을 참조하세요.
  - 지리 공간 함수에 대한 지원이 추가되었습니다.
  - 조인의 디스크로 유출에 대한 지원이 추가되었습니다.
  - Redshift 커넥터에 대한 지원이 추가되었습니다.
- Spark
  - SPARK-20640을 백포트했습니다. 이를 통해 `spark.shuffle.registration.timeout` 및 `spark.shuffle.registration.maxAttempts` 속성을 사용하여 셔플 등록에 대한 제시도와 rpc 제한 시간을 구성할 수 있습니다.
  - 사용자 지정 OutputFormat를 비 HDFS 위치에 기록할 때 발생하는 오류를 교정하는 SPARK-21549를 백포트했습니다.
- 하둡-13270을 백포트했습니다.
- Numpy, Scipy 및 Matplotlib 라이브러리가 기본 Amazon EMR AMI에서 제거되었습니다. 애플리케이션에서 이러한 라이브러리가 필요한 경우, 애플리케이션 리포지토리에서 라이브러리를 사용할 수 있기 때문에 부트스트랩 작업 시 `yum install`을 사용하여 모든 노드에 라이브러리를 설치할 수 있습니다.
- Amazon EMR 기본 AMI는 더 이상 애플리케이션 RPM 패키지를 포함하고 있지 않기 때문에 RPM 패키지가 더 이상 클러스터 노드에 존재하지 않습니다. 이제 사용자 지정 AMI 및 Amazon EMR 기본 AMI는 Amazon S3에서 RPM 패키지 리포지토리를 참조합니다.

- Amazon EC2에 초당 요금이 도입됨에 따라 이제 기본 스케일 다운 동작은 인스턴스 시간에 종료가 아닌 작업 완료 시 종료입니다. 자세한 내용은 [클러스터 스케일 다운 구성](#)을 참조하세요.

알려진 문제

- MXNet에는 OpenCV 라이브러리가 포함되어 있지 않습니다.
- Hive 2.3.1에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 [Statistics in Hive](#)를 참조하십시오.

릴리스 5.9.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.9.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.8.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 10월 5일

최근 기능 업데이트: 2017년 10월 12일

업그레이드

이 릴리스에서는 다음 버전을 포함하도록 다음 애플리케이션과 구성 요소를 업그레이드했습니다.

- AWS SDK for Java 버전 1.11.183
- Flink 1.3.2
- Hue 4.0.1
- Pig 0.17.0
- Presto 0.184

새로운 특성

- Livy 지원이 추가되었습니다(버전 0.4.0-incubating). 자세한 내용은 Apache Livy 단원을 참조하십시오.
- Spark용 Hue Notebook에 대한 지원이 추가되었습니다.
- i3 시리즈 Amazon EC2 인스턴스에 대한 지원이 추가되었습니다(2017년 10월 12일).

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Spark가 수동 크기 조절 또는 자동 조정 정책 요청으로 인한 노드 종료를 보다 원활하게 처리할 수 있도록 도움을 주는 새로운 기능 세트를 추가했습니다. 자세한 내용은 노드 서비스 해제 동작 구성 단원을 참조하십시오.
  - SSL은 블록 전송 서비스의 전송 중 암호화 시 3DES를 대신해 사용되며, AES-NI와 함께 Amazon EC2 인스턴스 유형을 사용할 때 성능을 개선합니다.
  - SPARK-21494를 백포트했습니다.
- Zeppelin
  - ZEPPELIN-2377을 백포트했습니다.
- HBase
  - HBase-18533 패치를 추가했습니다. 이 패치는 `hbase-site` 구성 분류를 사용하여 HBase BucketCache 구성을 위한 추가적인 값들을 허용합니다.
- Hue
  - Hue의 Hive 쿼리 편집기에 대한 AWS Glue 데이터 카탈로그 지원이 추가되었습니다.
  - 기본적으로 Hue의 수퍼유저는 Amazon EMR IAM 역할에서 액세스할 수 있는 모든 파일에 액세스할 수 있습니다. 새로 생성된 사용자에게는 Amazon S3 파일 브라우저에 액세스할 수 있는 권한이 자동으로 부여되지 않기 때문에 사용자의 그룹에 대해 활성화된 `filebrowser.s3_access` 권한이 있어야 합니다.
- AWS Glue 데이터 카탈로그를 사용하여 생성된 기본 JSON 데이터를 액세스할 수 없는 문제가 해결되었습니다.

알려진 문제

- 모든 애플리케이션을 설치하고 기본 Amazon EBS 루트 볼륨 크기를 변경하지 않은 상태에서 클러스터가 시작되지 않습니다. 해결 방법으로의 `aws emr create-cluster` 명령을 사용하고 더 큰 `--ebs-root-volume-size` 파라미터를 AWS CLI 지정합니다.
- Hive 2.3.0에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 [Statistics in Hive](https://cwiki.apache.org/confluence/display/Hive/StatsDev)를 참조하십시오.

릴리스 5.8.2

다음 릴리스 정보에는 Amazon EMR 릴리스 5.8.2에 대한 정보가 포함됩니다. 변경 사항은 5.8.1에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

릴리스 5.8.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.8.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.8.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

릴리스 5.8.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.8.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.7.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2017년 8월 10일

최근 기능 업데이트: 2017년 9월 25일

업그레이드

이 릴리스에서는 다음 버전을 포함하도록 다음 애플리케이션과 구성 요소를 업그레이드했습니다.

- AWS SDK 1.11.160
- Flink 1.3.1
- Hive 2.3.0 자세한 내용은 Apache Hive 사이트에서 Release notes를 참조하세요.
- Spark 2.2.0 자세한 내용은 Apache Spark 사이트에서 Release notes를 참조하세요.

새로운 특성

- 애플리케이션 이력 보기 지원이 추가되었습니다(2017년 9월 25일). 자세한 내용은 Amazon EMR 관리 안내서에서 애플리케이션 기록 보기를 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- AWS Glue 데이터 카탈로그와 통합
  - Hive 및 Spark SQL이 AWS Glue 데이터 카탈로그를 Hive 메타데이터 스토어로 사용할 수 있는 기능이 추가되었습니다. 자세한 내용은 Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용 및 Amazon EMR에서 Spark와 함께 AWS Glue 데이터 카탈로그 사용 섹션을 참조하세요.
- YARN 애플리케이션에 대한 기록 데이터 및 Spark 애플리케이션에 대한 추가 세부 정보를 볼 수 있는 클러스터 정보에 애플리케이션 이력을 추가했습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 애플리케이션 기록 보기를 참조하세요.
- Oozie
  - OOZIE-2748을 백포트했습니다.
- Hue
  - HUE-5859를 백포트했습니다

- HBase
  - `getMasterInitializedTime`을 사용하여 JMX(Java Management Extensions)를 통해 HBase 마스터 서버 시작 시간을 공개하도록 패치를 추가했습니다.
  - 클러스터 시작 시간을 개선하는 패치를 추가했습니다.

알려진 문제

- 모든 애플리케이션을 설치하고 기본 Amazon EBS 루트 볼륨 크기를 변경하지 않은 상태에서 클러스터가 시작되지 않습니다. 해결 방법으로의 `aws emr create-cluster` 명령을 사용하고 더 큰 `--ebs-root-volume-size` 파라미터를 AWS CLI 지정합니다.
- Hive 2.3.0에서는 기본적으로 `hive.compute.query.using.stats=true`를 설정합니다. 따라서 쿼리에서는 데이터를 직접 가져오지 않고 기존 통계에서 데이터를 가져오므로 혼동될 수 있습니다. 예를 들어, `hive.compute.query.using.stats=true`인 테이블이 있을 때 새 파일을 `LOCATION` 테이블에 업로드하고 테이블에서 `SELECT COUNT(*)` 쿼리를 실행하면 추가된 행이 선택되지 않고 통계에 있는 수치가 반환됩니다.

  차선책으로 `ANALYZE TABLE` 명령을 사용하여 새 통계를 수집하거나 `hive.compute.query.using.stats=false`를 설정합니다. 자세한 내용은 Apache Hive 설명서의 Statistics in Hive를 참조하십시오.
- Spark - Spark를 사용할 때, apppusher 대몬(daemon)과 관련된 파일 핸들러 유출 문제가 있습니다. 이는 몇 시간 또는 며칠 후에 장시간 실행되는 Spark 작업에 나타날 수 있습니다. 이 문제를 해결하려면 마스터 노드에 연결하고 `sudo /etc/init.d/apppusher stop`을 입력합니다. 이렇게 하면 apppusher 대몬(daemon)이 중단되며 Amazon EMR이 이를 자동으로 다시 시작합니다.
- 애플리케이션 이력
  - 작동 중단 Spark 실행기의 기록 데이터는 사용할 수 없습니다.
  - 진행 중인 암호화를 사용하기 위해 보안 구성을 사용하는 클러스터에서는 애플리케이션 이력을 사용할 수 없습니다.

릴리스 5.7.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.7.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.6.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 7월 13일

업그레이드

- Flink 1.3.0
- Phoenix 4.11.0
- Zeppelin 0.7.2

새로운 특성

- 클러스터를 생성할 때 사용자 지정 Amazon Linux AMI를 지정하는 기능을 추가했습니다. 자세한 내용은 사용자 지정 AMI 사용을 참조하세요.

변경 사항, 향상된 기능 및 해결된 문제

- HBase
  - HBase 읽기 전용 복제본 클러스터를 구성하는 기능을 추가했습니다. 읽기 전용 복제본 클러스터 사용을 참조하세요.
  - 다양한 버그 수정 사항 및 기능 향상
- Presto - `node.properties` 구성 기능을 추가했습니다.
- YARN - `container-log4j.properties` 구성 기능을 추가했습니다.
- Sqoop - SQOOP-2880을 백포트하여 Sqoop 임시 디렉터리를 설정할 수 있는 인수를 도입했습니다.

릴리스 5.6.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.6.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.5.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 6월 5일

업그레이드

- Flink 1.2.1
- HBase 1.3.1
- Mahout 0.13.0. Amazon EMR 버전 5.0 이상에서 Spark 2.x를 지원하는 첫 번째 버전의 Mahout입니다.
- Spark 2.1.1

변경 사항, 향상된 기능 및 해결된 문제

- Presto
  - 보안 구성에서 전송 데이터 암호화를 사용하여 Presto 노드 간 SSL/TLS 보안 통신을 활성화할 수 있는 기능이 추가되었습니다. 자세한 내용은 전송 중인 데이터 암호화를 참조하세요.
  - Presto 7661의 백포트와 함께 `VERBOSE` 옵션이 `EXPLAIN ANALYZE` 문에 추가되면서 쿼리 계획에 대해 더욱 자세한 저레벨 통계를 보고할 수 있게 되었습니다.

릴리스 5.5.3

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.3에 대한 정보가 포함됩니다. 변경 사항은 5.5.2에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

릴리스 5.5.2

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.2에 대한 정보가 포함됩니다. 변경 사항은 5.5.1에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

릴리스 5.5.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.5.0 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

릴리스 5.5.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.5.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.4.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 4월 26일

업그레이드

- Hue 3.12
- Presto 0.170
- Zeppelin 0.7.1
- ZooKeeper 3.4.10

변경 사항, 향상된 기능 및 해결된 문제

- Spark
  - Spark Patch (SPARK-20115) fix DAGScheduler to recompute all the lost shuffle blocks when external shuffle service is unavailable이 Spark의 2.1.0 버전으로 백포트되었습니다(이 패치는 이 릴리스에 포함됨).
- Flink
  - Flink는 이제 Scala 2.11로 빌드됩니다. 따라서 Scala API 및 라이브러리를 사용하는 경우에는 프로젝트에 Scala 2.11을 사용하는 것이 좋습니다.
  - `HADOOP_CONF_DIR` 및 `YARN_CONF_DIR` 기본값이 올바로 설정되지 않아서 `start-scala-shell.sh` 작업이 중단되던 문제가 해결되었습니다. 또한 `env.hadoop.conf.dir` 또는 `env.yarn.conf.dir` 구성 분류에서 `/etc/flink/conf/flink-conf.yaml` 및 `flink-conf`을 사용하여 이 값들을 설정할 수 있는 기능도 추가되었습니다.
  - 새로운 EMR 전용 명령인 `flink-scala-shell`이 `start-scala-shell.sh` 래퍼로 도입되었습니다. 앞으로 `start-scala-shell`이 아닌 이 명령을 사용하는 것이 좋습니다. 새로운 명령은

실행이 간편합니다. 예를 들어 `flink-scala-shell -n 2`는 Flink Scala 셸을 작업 병렬 처리 2로 시작합니다.

  - 새로운 EMR 전용 명령인 `flink-yarn-session`이 `yarn-session.sh` 래퍼로 도입되었습니다. 앞으로 `yarn-session`이 아닌 이 명령을 사용하는 것이 좋습니다. 새로운 명령은 실행이 간편합니다. 예를 들어, `flink-yarn-session -d -n 2`는 연결이 끊어진 상태에서 태스크 관리자 2개를 사용하여 장기 실행 Flink 세션을 시작합니다.
  - (FLINK-6125) commons httpclient is not shaded anymore in Flink 1.2 문제를 해결했습니다
- Presto
  - LDAP 인증에 대한 지원이 추가되었습니다. 단, Amazon EMR 기반 Presto에서 LDAP를 사용하려면 Presto 코디네이터의 HTTPS 액세스를 활성화해야 합니다(`config.properties`의 `http-server.https.enabled=true`). 구성에 대한 자세한 내용은 Presto 설명서에서 LDAP authentication을 참조하세요.
  - `SHOW GRANTS` 지원이 추가되었습니다.
- Amazon EMR 기본 Linux AMI
  - Amazon EMR 릴리스는 이제 Amazon Linux 2017.03을 기반으로 합니다. 자세한 내용은 Amazon Linux AMI 2017.03 릴리스 정보를 참조하세요.
  - Amazon EMR 기본 Linux 이미지에서 Python 2.6을 제거했습니다. 이제는 Python 2.7 및 3.4가 기본 설치됩니다. 필요하다면 Python 2.6을 수동으로 설치할 수도 있습니다.

릴리스 5.4.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.4.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.3.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 3월 8일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Flink 1.2.0로 업그레이드
- Hbase 1.3.0로 업그레이드
- Phoenix 4.9.0로 업그레이드

> **Note**
>
> 이전 버전의 Amazon EMR에서 Amazon EMR 버전 5.4.0 이상으로 업그레이드하고 보조 인덱싱을 사용하는 경우 [Apache Phoenix 설명서](#)에 설명된 대로 로컬 인덱스를 업그레이드합니다. Amazon EMR은 `hbase-site` 분류에서 필수 구성을 제거하지만 인덱스를 다시 채워야 합니다. 인덱스의 온라인 및 오프라인 업그레이드가 지원됩니다. 온라인 업그레이드는 기본값이므로, 버전 4.8.0 이상인 Phoenix 클라이언트에서 초기화하는 동안 인덱스가 다시 채워집니다. 오프라인 업그레이드를 지정하려면 `phoenix.client.localIndexUpgrade` 분류에서 `phoenix-site` 구성을 false로 설정한 다음 SSH를 마스터 노드로 설정하여 `psql [zookeeper] -1`을 실행합니다.

- Presto 0.166로 업그레이드
- Zeppelin 0.7.0로 업그레이드

변경 사항 및 기능 향상

릴리스 레이블 emr-5.4.0의 Amazon EMR 릴리스에 대한 변경 사항은 다음과 같습니다.

- r4 인스턴스에 대한 지원을 추가했습니다. [Amazon EC2 인스턴스 유형](#)을 참조하세요.

릴리스 5.3.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.3.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.3.0 릴리스에 관련됩니다.

릴리스 날짜: 2017년 2월 7일

Zeppelin 패치를 백포트하고 Amazon EMR의 기본 AMI를 업데이트하는 부수적 변경 내용을 포함합니다.

릴리스 5.3.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.3.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.2.1 릴리스에 관련됩니다.

릴리스 날짜: 2017년 1월 26일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Hive 2.1.1으로 업그레이드
- Hue 3.11.0로 업그레이드
- Spark 2.1.0로 업그레이드
- Oozie 4.3.0로 업그레이드
- Flink 1.1.4로 업그레이드

변경 사항 및 기능 향상

릴리스 레이블 emr-5.3.0의 Amazon EMR 릴리스에 대한 변경 사항은 다음과 같습니다.

- `interpreters_shown_on_wheel` 파일에 지정된 순서와 상관없이 `hue.ini` 설정을 사용하여 노트북 선택 휠에 첫 번째로 표시할 인터프리터를 구성할 수 있는 Hue 패치를 추가했습니다.
- Hive의 `hive-parquet-logging` 파일에서 값을 구성하는 데 사용할 수 있는 `parquet-logging.properties` 구성 분류를 추가했습니다.

릴리스 5.2.2

다음 릴리스 정보에는 Amazon EMR 릴리스 5.2.2에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.2.1 릴리스에 관련됩니다.

릴리스 날짜: 2017년 5월 2일

이전 릴리스에서 해결된 알려진 문제

- SPARK-194459의 백포트와 함께 char/varchar 열이 포함된 ORC 테이블에서 데이터를 읽어오지 못하던 문제가 해결되었습니다.

릴리스 5.2.1

다음 릴리스 정보에는 Amazon EMR 릴리스 5.2.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.2.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 12월 29일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Presto 0.157.1로 업그레이드되었습니다. 자세한 내용은 Presto 설명서에서 Presto 릴리스 정보를 참조하세요.
- Zookeeper 3.4.9로 업그레이드되었습니다. 자세한 내용은 Apache ZooKeeper 설명서에서 ZooKeeper 릴리스 정보를 참조하세요.

변경 사항 및 기능 향상

릴리스 레이블 emr-5.2.1의 Amazon EMR 릴리스에 대한 변경 사항은 다음과 같습니다.

- Amazon EMR 버전 4.8.3 이상(5.0.0, 5.0.3 및 5.2.0 제외)에서 Amazon EC2 m4.16xlarge 인스턴스 유형에 대한 지원을 추가했습니다.
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.
- 이제 Flink 및 YARN 구성 경로가 `/etc/default/flink`에서 기본적으로 설정되므로 `FLINK_CONF_DIR` 또는 `HADOOP_CONF_DIR` 드라이버 스크립트를 실행하여 Flink 작업을 시작할 때 환경 변수 `flink` 및 `yarn-session.sh`을 설정할 필요가 없습니다.
- FlinkKinesisConsumer 클래스에 대한 지원을 추가했습니다.

이전 릴리스에서 해결된 알려진 문제

- 대규모 클러스터에서 동일한 파일의 복제 및 삭제 간 경합 때문에 ReplicationMonitor 스레드가 장시간 정체될 수 있는 하둡 문제를 수정했습니다.
- 작업 상태가 성공적으로 업데이트되지 않은 경우 ControlledJob#toString이 null 포인터 예외(NPE)와 함께 실패하는 문제를 수정했습니다.

릴리스 5.2.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.2.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.1.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 11월 21일

변경 사항 및 기능 향상

이 릴리스에서는 다음과 같은 변경 사항 및 기능 향상을 사용할 수 있습니다.

- HBase에 대한 Amazon S3 스토리지 모드를 추가했습니다.
- HBase rootdir의 Amazon S3 위치를 지정할 수 있습니다. 자세한 내용은 Amazon S3용 HBase를 참조하세요.

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Spark 2.0.2로 업그레이드

이전 릴리스에서 해결된 알려진 문제

- EBS 전용 인스턴스 유형에서 /mnt가 2TB로 제약되는 문제를 수정했습니다.
- instance-controller 및 logpusher 로그가 매시간 회전하는 정상적인 og4j-configured .log 파일 대신 해당 .out 파일로 출력되는 문제를 수정했습니다. .out 파일은 회전하지 않으므로 결국 /emr 파티션이 파일로 가득 찰 수 있습니다. 이 문제는 하드웨어 가상 머신(HVM) 인스턴스 유형에만 영향을 미칩니다.

릴리스 5.1.0

다음 릴리스 정보에는 Amazon EMR 릴리스 5.1.0에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.0.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 11월 3일

변경 사항 및 기능 향상

이 릴리스에서는 다음과 같은 변경 사항 및 기능 향상을 사용할 수 있습니다.

- Flink 1.1.3에 대한 지원을 추가했습니다.
- Presto는 Hue의 노트북 섹션에서 옵션으로 추가되었습니다.

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- HBase 1.2.3로 업그레이드
- Zeppelin 0.6.2로 업그레이드

이전 릴리스에서 해결된 알려진 문제

- ORC 파일이 이전 Amazon EMR 4.x 버전만큼 좋은 성능을 나타내지 않는 Amazon S3의 Tez 쿼리 문제를 수정했습니다.

릴리스 5.0.3

다음 릴리스 정보에는 Amazon EMR 릴리스 5.0.3에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 5.0.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 10월 24일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Hadoop 2.7.3로 업그레이드
- Presto 0.152.3로 업그레이드했습니다. 여기에는 Presto 웹 인터페이스에 대한 지원이 포함됩니다. 포트 8889를 사용하여 Presto 코디네이터의 Presto 웹 인터페이스에 액세스할 수 있습니다. Presto 웹 인터페이스에 대한 자세한 내용은 Presto 설명서에서 Web interface를 참조하세요.
- Spark 2.0.1으로 업그레이드
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.

릴리스 5.0.0

릴리스 날짜: 2016년 7월 27일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Hive 2.1으로 업그레이드
- Presto 0.150로 업그레이드
- Spark 2.0로 업그레이드
- Hue 3.10.0로 업그레이드
- Pig 0.16.0로 업그레이드
- Tez 0.8.4로 업그레이드
- Zeppelin 0.6.1으로 업그레이드

변경 사항 및 기능 향상

릴리스 레이블 emr-5.0.0 이상의 Amazon EMR 릴리스에 대한 변경 사항은 다음과 같습니다.

- Amazon EMR은 최신 오픈 소스 버전의 Hive(버전 2.1)와 Pig(버전 0.16.0)를 지원합니다. 이전에 Amazon EMR에서 Hive나 Pig를 사용한 적이 있는 경우 일부 사용 사례에 영향을 미칠 수 있습니다. 자세한 내용은 Hive 및 Pig를 참조하십시오.
- Hive 및 Pig용 기본 실행 엔진은 이제 Tez입니다. 이 항목을 변경하려면 각각 `hive-site` 및 `pig-properties` 구성 분류에서 해당 값을 편집합니다.
- 향상된 단계 디버깅 기능이 추가되었습니다. 서비스에서 원인을 결정할 수 있는 경우 이 기능을 사용하여 단계 실패의 근본 원인을 확인할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 향상된 단계 디버깅을 참조하세요.
- 이전에 "-Sandbox"로 끝났던 애플리케이션은 더 이상 해당 접미사를 사용하지 않습니다. 이로 인해 예를 들어 스크립트를 사용하여 이 애플리케이션이 있는 클러스터를 시작하는 경우 자동화가 중단될 수 있습니다. 다음 테이블에서는 Amazon EMR 5.0.0 대비 Amazon EMR 4.7.2의 애플리케이션 이름을 보여줍니다.

  애플리케이션 이름 변경

Amazon EMR 4.7.2	Amazon EMR 5.0.0
Oozie-Sandbox	Oozie
Presto-Sandbox	Presto
Sqoop-Sandbox	Sqoop
Zeppelin-Sandbox	Zeppelin

Amazon EMR 4.7.2	Amazon EMR 5.0.0
ZooKeeper-Sandbox	ZooKeeper

- Spark는 이제 Scala 2.11용으로 컴파일됩니다.
- 이제 기본 JVM은 Java 8입니다. 모든 애플리케이션은 Java 8 실행 시간을 사용하여 실행됩니다. 애플리케이션의 바이트 코드 대상에는 변경 사항이 없습니다. 대부분의 애플리케이션은 계속해서 Java 7을 대상으로 합니다.
- Zeppelin에는 이제 인증 기능이 포함됩니다. 자세한 내용은 Zeppelin을 참조하십시오.
- 암호화 옵션을 더 쉽게 생성하고 적용할 수 있도록 보안 구성에 대한 지원을 추가했습니다. 자세한 내용은 데이터 암호화를 참조하세요.

릴리스 4.9.5

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.5에 대한 정보가 포함됩니다. 변경 사항은 4.9.4에 관련됩니다.

최초 릴리스 날짜: 2018년 8월 29일

변경 사항, 향상된 기능 및 해결된 문제

- HBase
  - 이 릴리스는 잠재적인 보안 취약점을 해결합니다.

릴리스 4.9.4

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.4에 대한 정보가 포함됩니다. 변경 사항은 4.9.3에 관련됩니다.

최초 릴리스 날짜: 2018년 3월 29일

변경 사항, 향상된 기능 및 해결된 문제

- 잠재적인 취약성을 해결하기 위해 Amazon EMR에 대한 기본 Amazon Linux AMI의 Amazon Linux 커널을 업데이트했습니다.

릴리스 4.9.3

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.3에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.9.2 릴리스에 관련됩니다.

최초 릴리스 날짜: 2018년 1월 22일

변경 사항, 향상된 기능 및 해결된 문제

- 추론적 실행과 관련된 취약성을 해결하기 위해 Amazon EMR의 기본 Amazon Linux AMI에서 Amazon Linux 커널을 업데이트했습니다(CVE-2017-5715, CVE-2017-5753, CVE-2017-5754). 자세한 내용은 https://aws.amazon.com/security/security-bulletins/AWS-2018-013/ 단원을 참조하십시오.

릴리스 4.9.2

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.2에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.9.1 릴리스에 관련됩니다.

릴리스 날짜: 2017년 7월 13일

이번 릴리스에서 일반적인 변경, 버그 수정 및 기능 개선이 이루어졌습니다.

릴리스 4.9.1

다음 릴리스 정보에는 Amazon EMR 릴리스 4.9.1에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.4 릴리스에 관련됩니다.

릴리스 날짜: 2017년 4월 10일

이전 릴리스에서 해결된 알려진 문제

- HIVE-9976 및 HIVE-10106 백포트
- YARN에서 다수의 노드(2,000개 이상) 및 컨테이너(5,000개 이상)가 메모리 부족 오류를 야기하던 문제(예: `"Exception in thread 'main' java.lang.OutOfMemoryError"`)가 수정되었습니다.

변경 사항 및 기능 향상

릴리스 레이블 emr-4.9.1의 Amazon EMR 릴리스에 대한 변경 사항은 다음과 같습니다.

- Amazon EMR 릴리스는 이제 Amazon Linux 2017.03을 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2017.03-release-notes/ 단원을 참조하십시오.
- Amazon EMR 기본 Linux 이미지에서 Python 2.6을 제거했습니다. 필요하다면 Python 2.6을 수동으로 설치할 수도 있습니다.

릴리스 4.8.4

다음 릴리스 정보에는 Amazon EMR 릴리스 4.8.4에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.3 릴리스에 관련됩니다.

릴리스 날짜: 2017년 2월 7일

이번 릴리스에서 일반적인 변경, 버그 수정 및 기능 개선이 이루어졌습니다.

릴리스 4.8.3

다음 릴리스 정보에는 Amazon EMR 릴리스 4.8.3에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.2 릴리스에 관련됩니다.

릴리스 날짜: 2016년 12월 29일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Presto 0.157.1로 업그레이드되었습니다. 자세한 내용은 Presto 설명서에서 Presto 릴리스 정보를 참조하세요.
- Spark 1.6.3으로 업그레이드되었습니다. 자세한 내용은 Apache Spark 설명서에서 Spark 릴리스 정보를 참조하세요.
- ZooKeeper 3.4.9로 업그레이드되었습니다. 자세한 내용은 Apache ZooKeeper 설명서에서 ZooKeeper 릴리스 정보를 참조하세요.

변경 사항 및 기능 향상

릴리스 레이블 emr-4.8.3의 Amazon EMR 릴리스에 대한 변경 사항은 다음과 같습니다.

- Amazon EMR 버전 4.8.3 이상(5.0.0, 5.0.3 및 5.2.0 제외)에서 Amazon EC2 m4.16xlarge 인스턴스 유형에 대한 지원을 추가했습니다.

- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.

이전 릴리스에서 해결된 알려진 문제

- 대규모 클러스터에서 동일한 파일의 복제 및 삭제 간 경합 때문에 ReplicationMonitor 스레드가 장시간 정체될 수 있는 하둡 문제를 수정했습니다.
- 작업 상태가 성공적으로 업데이트되지 않은 경우 ControlledJob#toString이 null 포인터 예외(NPE)와 함께 실패하는 문제를 수정했습니다.

릴리스 4.8.2

다음 릴리스 정보에는 Amazon EMR 릴리스 4.8.2에 대한 정보가 포함됩니다. 변경 사항은 Amazon EMR 4.8.0 릴리스에 관련됩니다.

릴리스 날짜: 2016년 10월 24일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- Hadoop 2.7.3로 업그레이드
- Presto 0.152.3로 업그레이드했습니다. 여기에는 Presto 웹 인터페이스에 대한 지원이 포함됩니다. 포트 8889를 사용하여 Presto 코디네이터의 Presto 웹 인터페이스에 액세스할 수 있습니다. Presto 웹 인터페이스에 대한 자세한 내용은 Presto 설명서에서 Web interface를 참조하세요.
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.09를 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.09-release-notes/ 단원을 참조하십시오.

릴리스 4.8.0

릴리스 날짜: 2016년 9월 7일

업그레이드

이 릴리스에서는 다음 업그레이드를 사용할 수 있습니다.

- HBase 1.2.2로 업그레이드

- Presto-Sandbox 0.151으로 업그레이드
- Tez 0.8.4로 업그레이드
- Zeppelin-Sandbox 0.6.1으로 업그레이드

변경 사항 및 기능 향상

릴리스 레이블 emr-4.8.0의 Amazon EMR 릴리스에 대한 변경 사항은 다음과 같습니다.

- YARN에서 인스턴스가 종료되었기 때문에 더 이상 존재하지 않는 컨테이너를 ApplicationMaster가 정리하려고 시도하는 문제를 수정했습니다.
- Oozie 예제에서 Hive2 작업에 대한 hive-server2 URL을 수정했습니다.
- 추가 Presto 카탈로그에 대한 지원을 추가했습니다.
- 백포트된 패치: HIVE-8948, HIVE-12679, HIVE-13405, PHOENIX-3116, HADOOP-12689
- 암호화 옵션을 더 쉽게 생성하고 적용할 수 있도록 보안 구성에 대한 지원을 추가했습니다. 자세한 내용은 데이터 암호화를 참조하세요.

릴리스 4.7.2

다음 릴리스 정보에는 Amazon EMR 릴리스 4.7.2에 대한 정보가 포함됩니다.

릴리스 날짜: 2016년 7월 15일

Features

이 릴리스에서는 다음 기능을 사용할 수 있습니다.

- Mahout 0.12.2로 업그레이드
- Presto 0.148으로 업그레이드
- Spark 1.6.2로 업그레이드
- 이제 URI를 파라미터로 사용하여 EMRFS과 함께 사용할 AWSCredentialsProvider를 생성할 수 있습니다. 자세한 내용은 EMRFS에 대한 AWSCredentialsProvider 생성을 참조하십시오.
- 이제 EMRFS에서 사용자는 `emrfs-site.xml`의 `fs.s3.consistent.dynamodb.endpoint` 속성을 사용하여 일관된 보기 메타데이터에 대한 사용자 지정 DynamoDB 엔드포인트를 구성할 수 있습니다.
- `/usr/bin`이라고 하는 `spark-example`에 스크립트가 추가되면서 `/usr/lib/spark/spark/bin/run-example`을 래핑하여 직접 예제를 실행할 수 있습니다. 예를 들어, Spark 배포와 함께 제

공되는 SparkPi 예제를 실행하려면 명령줄에서 또는 `spark-example SparkPi 100`을 API에서 단계로 사용하여 `command-runner.jar`를 실행할 수 있습니다.

이전 릴리스에서 해결된 알려진 문제

- Oozie 와 함께 Spark도 설치된 경우 `spark-assembly.jar`이 올바른 위치에 있지 않아서 Spark 애플리케이션이 Oozie와 함께 시작되지 못하는 문제를 수정했습니다.
- YARN 컨테이너에서 Spark Log4j 기반 로깅 문제를 수정했습니다.

릴리스 4.7.1

릴리스 날짜: 2016년 6월 10일

이전 릴리스에서 해결된 알려진 문제

- VPC에서 프라이빗 서브넷을 사용하여 시작된 클러스터의 스타트업 시간을 연장하는 문제를 수정했습니다. 이 버그는 Amazon EMR 4.7.0 릴리스를 사용하여 시작된 클러스터에만 영향을 미칩니다.
- Amazon EMR 4.7.0 릴리스를 사용하여 시작된 클러스터의 경우 Amazon EMR에서 파일 목록을 부적절하게 처리하는 문제를 수정했습니다.

릴리스 4.7.0

> **Important**
>
> Amazon EMR 4.7.0은 더 이상 사용되지 않습니다. 대신 Amazon EMR 4.7.1 이상을 사용합니다.

릴리스 날짜: 2016년 6월 2일

Features

이 릴리스에서는 다음 기능을 사용할 수 있습니다.

- Apache Phoenix 4.7.0 추가
- Apache Tez 0.8.3 추가

- HBase 1.2.1으로 업그레이드
- Mahout 0.12.0로 업그레이드
- Presto 0.147으로 업그레이드
- 를 1.10.75 AWS SDK for Java 로 업그레이드
- 사용자가 로컬 모드에서 Pig를 실행할 수 있도록 `mapreduce.cluster.local.dir`의 `mapred-site.xml` 속성에서 최종 플래그가 제거되었습니다.

클러스터에서 사용할 수 있는 Amazon Redshift JDBC 드라이버

이제 Amazon Redshift JDBC 드라이버는 `/usr/share/aws/redshift/jdbc`에 포함됩니다. `/usr/share/aws/redshift/jdbc/RedshiftJDBC41.jar`은 JDBC 4.1 호환 Amazon Redshift 드라이버이고 `/usr/share/aws/redshift/jdbc/RedshiftJDBC4.jar`은 JDBC 4.0 호환 Amazon Redshift 드라이버입니다. 자세한 내용은 Amazon Redshift 관리 안내서에서 [JDBC 연결 구성]을 참조하세요.

Java 8

Presto를 제외하고 모든 애플리케이션에 사용되는 기본 JDK는 OpenJDK 1.7입니다. 하지만 OpenJDK 1.7 및 1.8이 모두 설치됩니다. 애플리케이션에 대해 `JAVA_HOME`을 설정하는 방법에 대한 자세한 내용은 [Java 8을 사용하도록 애플리케이션 구성]을 참조하세요.

이전 릴리스에서 해결된 알려진 문제

- emr-4.6.0에서 Amazon EMR에 사용되는 처리량에 최적화된 HDD(st1) EBS 볼륨의 성능에 상당한 영향을 미치는 커널 문제를 수정했습니다.
- 하둡을 애플리케이션으로 선택하지 않고 HDFS 암호화 영역을 지정할 경우 클러스터가 실패하는 문제를 수정했습니다.
- 기본 HDFS 쓰기 정책을 `RoundRobin`에서 `AvailableSpaceVolumeChoosingPolicy`로 변경했습니다. RoundRobin 구성에서 일부 볼륨이 제대로 이용되지 않았으며 결과적으로 코어 노드가 실패하고 HDFS가 불안정한 상태가 되었습니다.
- 일관된 보기에 대한 기본 DynamoDB 메타데이터 테이블을 생성할 때 예외가 발생하는 EMRFS CLI 문제를 수정했습니다.
- 멀티파트 이름 바꾸기 및 복사 작업 중에 발생할 수 있는 EMRFS의 교착 문제를 수정했습니다.
- CopyPart 크기가 기본적으로 5MB로 설정되는 EMRFS 문제를 수정했습니다. 기본값은 이제 128MB로 올바르게 설정됩니다.

- 서비스를 중단할 수 없는 Zeppelin upstart 구성 문제를 수정했습니다.
- `s3a://`이 해당 클래스 경로에 올바르게 로드되지 않기 때문에 `/usr/lib/hadoop/hadoop-aws.jar` URI 스키마를 사용할 수 없는 Spark 및 Zeppelin 문제를 수정했습니다.
- HUE-2484를 백포트했습니다.
- HBase 브라우저 샘플 문제를 수정하기 위해 Hue 3.9.0(JIRA가 존재하지 않음)의 커밋을 백포트했습니다.
- HIVE-9073를 백포트했습니다.

릴리스 4.6.0

릴리스 날짜: 2016년 4월 21일

Features

이 릴리스에서는 다음 기능을 사용할 수 있습니다.

- HBase 1.2.0 추가
- Zookeeper-Sandbox 3.4.8 추가
- Presto-Sandbox 0.143로 업그레이드
- Amazon EMR 릴리스는 이제 Amazon Linux 2016.03.0을 기반으로 합니다. 자세한 내용은 https://aws.amazon.com/amazon-linux-ami/2016.03-release-notes/ 단원을 참조하십시오.

처리량에 최적화된 HDD(st1) EBS 볼륨 유형에 영향을 미치는 문제

Linux 커널 버전 4.2 이상의 문제는 EMR에 사용되는 처리량에 최적화된 HDD(st1) EBS 볼륨의 성능에 상당한 영향을 미칩니다. 이 릴리스(emr-4.6.0)는 커널 버전 4.4.5를 사용하므로 영향을 받습니다. 따라서 st1 EBS 볼륨을 사용하려는 경우 emr-4.6.0를 사용하지 않는 것이 좋습니다. emr-4.5.0 이전의 Amazon EMR 릴리스를 st1과 함께 사용하면 영향을 받지 않을 수 있습니다. 또한 향후 릴리스에서 수정을 제공합니다.

Python 기본값

이제 Python 3.4가 기본적으로 설치되지만, 시스템 기본값은 여전히 Python 2.7입니다. 부트스트랩 작업을 사용하여 Python 3.4를 시스템 기본값으로 구성할 수 있습니다. 구성 API를 사용하여 `/usr/bin/python3.4` 분류의 PYSPARK_PYTHON 내보내기를 `spark-env`로 설정하면 PySpark에서 사용되는 Python 버전에 영향을 미칠 수 있습니다.

Java 8

Presto를 제외하고 모든 애플리케이션에 사용되는 기본 JDK는 OpenJDK 1.7입니다. 하지만 OpenJDK 1.7 및 1.8이 모두 설치됩니다. 애플리케이션에 대해 `JAVA_HOME`을 설정하는 방법에 대한 자세한 내용은 [Java 8을 사용하도록 애플리케이션 구성](https://docs.aws.amazon.com/emr/latest/ReleaseGuide/emr-configure-apps.html#configuring-java8)을 참조하세요.

이전 릴리스에서 해결된 알려진 문제

- 생성된 암호로 인해 애플리케이션 프로비저닝이 가끔 임의로 실패하는 문제를 수정했습니다.
- 이전에는 `mysqld`가 모든 노드에 설치되었습니다. 이제 이 항목은 마스터 인스턴스에만 설치되며 선택한 애플리케이션에 `mysql-server`가 구성 요소로 포함되는 경우에만 설치됩니다. 현재 HCatalog, Hive, Hue, Presto-Sandbox, Sqoop-Sandbox 등의 애플리케이션에 `mysql-server` 구성 요소가 포함됩니다.
- `yarn.scheduler.maximum-allocation-vcores`를 기본값 32에서 80으로 변경했습니다. 이 변경에 따라 emr-4.4.0에 도입된 한 가지 문제가 수정됩니다. 이 문제는 코어 인스턴스 유형이 32 이상의 YARN vcore 세트를 갖춘 몇 가지 대규모 인스턴스 유형 중 하나인 클러스터에서 `maximizeResourceAllocation` 옵션을 사용하는 동안 주로 Spark에 발생합니다. 다시 말해서 c4.8xlarge, cc2.8xlarge, hs1.8xlarge, i2.8xlarge, m2.4xlarge, r3.8xlarge, d2.8xlarge 또는 m4.10xlarge가 이 문제의 영향을 받습니다.
- s3-dist-cp는 이제 모든 Amazon S3 추천에 EMRFS를 사용하며 더 이상 임시 HDFS 디렉터리로 준비하지 않습니다.
- 클라이언트 측 암호화 멀티파트 업로드에 대한 예외 처리 문제를 수정했습니다.
- 사용자가 Amazon S3 스토리지 클래스를 변경할 수 있는 옵션을 추가했습니다. 기본적으로 이 설정은 `STANDARD`입니다. `emrfs-site` 구성 분류 설정은 `fs.s3.storageClass`이며 가능한 값은 `STANDARD`, `STANDARD_IA` 및 `REDUCED_REDUNDANCY`입니다. 스토리지 클래스에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 [스토리지 클래스](https://docs.aws.amazon.com/AmazonS3/latest/dev/storage-class-intro.html)를 참조하세요.

릴리스 4.5.0

릴리스 날짜: 2016년 4월 4일

Features

이 릴리스에서는 다음 기능을 사용할 수 있습니다.

- Spark 1.6.1으로 업그레이드
- Hadoop 2.7.2로 업그레이드

- Presto 0.140로 업그레이드
- Amazon S3 서버 측 암호화에 대한 AWS KMS 지원이 추가되었습니다.

이전 릴리스에서 해결된 알려진 문제

- 노드를 재부팅한 후 MySQL 및 Apache 서버가 시작되지 않는 문제를 수정했습니다.
- Amazon S3에 파티션되지 않은 테이블이 저장되면 IMPORT가 올바르게 작동하지 않는 문제를 수정했습니다.
- Hive 테이블에 쓸 때 준비 디렉터리를 `/mnt/tmp` 대신 `/tmp`로 지정해야 하는 Presto 문제를 수정했습니다.

릴리스 4.4.0

릴리스 날짜: 2016년 3월 14일

Features

이 릴리스에서는 다음 기능을 사용할 수 있습니다.

- HCatalog 1.0.0 추가
- Sqoop-Sandbox 1.4.6 추가
- Presto 0.136로 업그레이드
- Zeppelin 0.5.6로 업그레이드
- Mahout 0.11.1으로 업그레이드
- `dynamicResourceAllocation`을 기본적으로 활성화했습니다.
- 릴리스에 대한 모든 구성 분류의 표를 추가했습니다. 자세한 내용은 애플리케이션 구성의 구성 분류 테이블을 참조하세요.

이전 릴리스에서 해결된 알려진 문제

- `maximizeResourceAllocation` 설정에서 YARN ApplicationMaster 데몬에 충분한 메모리가 예약되지 않는 문제를 수정했습니다.
- 사용자 지정 DNS에 발생하는 문제를 수정했습니다. 제공된 사용자 지정 항목 앞에 `resolve.conf`의 항목이 있으면 사용자 지정 항목을 확인할 수 없습니다. 이 동작은 기본 VPC 이

름 서버가 `resolve.conf`에 최상위 항목으로 삽입되는 VPC의 클러스터에서 영향을 받은 것입니다.

- 기본 Python이 버전 2.7으로 이동하고 해당 버전에 대해 boto가 설치되지 않는 문제를 수정했습니다.
- YARN 컨테이너와 Spark 애플리케이션이 고유의 Ganglia rrd(round robin database) 파일을 생성하여 인스턴스에 연결된 첫 번째 디스크가 가득 채워지는 문제를 수정했습니다. 이러한 수정으로 인해 YARN 컨테이너 수준 지표가 비활성화되었으며 Spark 애플리케이션 수준 지표가 비활성화되었습니다.
- 로그 푸셔에서 비어 있는 모든 로그 폴더가 삭제되는 문제를 수정했습니다. 이 문제의 영향으로 로그 푸셔가 `user`에서 비어 있는 `/var/log/hive` 폴더를 제거하기 때문에 Hive CLI가 로그를 작성할 수 없었습니다.
- Hive 가져오기에 영향을 미치는 문제를 수정했습니다. 이 문제는 파티셔닝에 영향을 미치며 결과적으로 가져오기 중 오류가 발생했습니다.
- EMRFS 및 s3-dist-cp에서 마침표가 포함된 버킷 이름이 올바르게 처리되지 않는 문제를 수정했습니다.
- 버전 관리를 사용하는 버킷에서 `_$folder$` 마커 파일이 연속적으로 생성되지 않도록 EMRFS의 동작을 변경했습니다. 이 변경은 버전 관리를 사용하는 버킷의 성능을 향상하는 데 도움이 될 수 있습니다.
- 클라이언트 측 암호화를 사용하는 경우를 제외하고 지침 파일을 사용하지 않도록 EMRFS의 동작을 변경했습니다. 클라이언트 측 암호화를 사용하는 동안 지침 파일을 삭제하려는 경우 emrfs-site.xml 속성인 `fs.s3.cse.cryptoStorageMode.deleteInstructionFiles.enabled`를 true로 설정할 수 있습니다.
- 2일 동인 집계 대상에서 로그를 유지하도록 Changed YARN 로그 집계를 변경했습니다. 기본 대상은 클러스터의 HDFS 스토리지입니다. 이 기간을 변경하려는 경우 클러스터를 생성할 때 `yarn.log-aggregation.retain-seconds` 분류를 사용하여 `yarn-site`의 값을 변경합니다. 평소와 같이, 클러스터를 생성할 때 `log-uri` 파라미터를 사용하여 애플리케이션 로그를 Amazon S3에 저장할 수 있습니다.

패치 적용됨

이 릴리스에는 다음과 같은 오픈 소스 프로젝트의 패치가 포함되었습니다.

- HIVE-9655
- HIVE-9183
- HADOOP-12810

릴리스 4.3.0

릴리스 날짜: 2016년 1월 19일

Features

이 릴리스에서는 다음 기능을 사용할 수 있습니다.

- Hadoop 2.7.1으로 업그레이드
- Spark 1.6.0로 업그레이드
- Ganglia를 3.7.2로 업그레이드
- Presto를 0.130로 업그레이드

Amazon EMR에서 `spark.dynamicAllocation.enabled`를 true로 설정한 경우 이 설정이 약간 변경되었습니다. 이 설정은 기본적으로 false입니다. 이 값을 true로 설정하면 이 설정은 `maximizeResourceAllocation` 설정을 통해 지정된 기본값에 영향을 미칩니다.

- `spark.dynamicAllocation.enabled`가 true로 설정되면 `spark.executor.instances`에서 `maximizeResourceAllocation`가 설정되지 않습니다.
- 이제 `spark.driver.memory` 설정은 `spark.executors.memory`가 설정되는 방식과 비슷한 방식으로 클러스터의 인스턴스 유형에 따라 구성됩니다. 하지만 Spark 드라이버 애플리케이션을 마스터 또는 코어 인스턴스 중 하나에서(예를 들면 각각 YARN 클라이언트 및 클러스터 모드에서) 실행할 수 있으므로 `spark.driver.memory` 설정이 이러한 두 인스턴스 그룹 중 더 작은 인스턴스 유형에 따라 설정됩니다.
- 이제 `spark.default.parallelism` 설정은 YARN 컨테이너에 사용할 수 있는 CPU 코어 수의 두 배로 설정됩니다. 이전 릴리스에서 이 설정은 해당 값의 절반이었습니다.
- Spark YARN 프로세스에서 예약된 메모리 오버헤드에 대한 계산이 더 정확하게 조정되어 Spark에 사용 가능한 총 메모리 양(즉, `spark.executor.memory`)이 약간 증가했습니다.

이전 릴리스에서 해결된 알려진 문제

- 이제 YARN 로그 집계가 기본적으로 활성화됩니다.
- YARN 로그 집계를 활성화하면 로그가 클러스터의 Amazon S3 로그 버킷에 푸시되지 않는 문제를 수정했습니다.
- 이제 YARN 컨테이너 크기의 새로운 최소값은 모든 노드 유형에서 32입니다.
- 대규모 클러스터의 마스터 노드에 과도한 디스크 I/O를 유발하는 Ganglia 문제를 수정했습니다.

- 클러스터를 종료할 때 애플리케이션 로그가 Amazon S3에 푸시되지 않는 문제를 수정했습니다.
- EMRFS CLI에서 특정 명령이 실패하는 문제를 수정했습니다.
- 종속성이 기본 SparkContext에 로드되지 않는 Zeppelin 문제를 수정했습니다.
- 크기 조정을 실행하여 인스턴스를 추가하려고 할 때 발생하는 문제를 수정했습니다.
- Hive에서 CREATE TABLE AS SELECT를 실행하면 Amazon S3에 대한 목록이 과도하게 직접 호출되는 문제를 수정했습니다.
- Hue, Oozie 및 Ganglia를 설치하면 대규모 클러스터가 제대로 프로비저닝되지 않는 문제를 수정했습니다.
- s3-dist-cp가 오류로 실패하더라도 0 종료 코드가 반환되는 문제를 수정했습니다.

패치 적용됨

이 릴리스에는 다음과 같은 오픈 소스 프로젝트의 패치가 포함되었습니다.

- OOZIE-2402
- HIVE-12502
- HIVE-10631
- HIVE-12213
- HIVE-10559
- HIVE-12715
- HIVE-10685

릴리스 4.2.0

릴리스 날짜: 2015년 11월 18일

Features

이 릴리스에서는 다음 기능을 사용할 수 있습니다.

- Ganglia 지원 추가
- Spark 1.5.2로 업그레이드
- Presto 0.125로 업그레이드
- Oozie를 4.2.0으로 업그레이드

- Zeppelin을 0.5.5로 업그레이드
- 를 1.10.27 AWS SDK for Java 로 업그레이드했습니다.

이전 릴리스에서 해결된 알려진 문제

- 기본 메타데이터 테이블 이름이 사용되지 않는 EMRFS CLI 문제를 수정했습니다.
- Amazon S3에서 ORC 지원 테이블을 사용할 때 발생하는 문제를 수정했습니다.
- Spark 구성에서 Python 버전이 일치하지 않는 경우 발생하는 문제를 수정했습니다.
- VPC에서 클러스터의 DNS 문제 때문에 YARN 노드 상태가 보고되지 않는 문제를 수정했습니다.
- YARN에서 노드를 폐기하면 애플리케이션이 정지되거나 새 애플리케이션을 예약할 수 없게 되는 문제를 수정했습니다.
- 클러스터가 TIMED_OUT_STARTING 상태로 종료될 때 발생하는 문제를 수정했습니다.
- EMRFS Scala 종속성을 다른 빌드에 포함시킬 때 발생하는 문제를 수정했습니다. Scala 종속성이 제거되었습니다.

애플리케이션 구성

애플리케이션에 대한 기본 구성을 재정의하기 위해 구성 객체를 제공할 수 있습니다. 간편 구문을 사용하여 구성을 제공하거나 JSON 파일에서 구성 객체를 참조할 수 있습니다. 구성 객체는 분류, 속성 및 선택적 중첩 구성으로 이루어져 있습니다. 속성은 변경할 애플리케이션 설정에 해당합니다. 단일 JSON 객체에서 여러 애플리케이션에 대해 다양한 분류를 지정할 수 있습니다.

> **Warning**
>
> Amazon EMR 설명 및 나열 API 작업은 사용자 지정 및 구성 가능한 설정을 생성하며, 이 설정은 일반 텍스트에서 Amazon EMR 작업 흐름의 일부로 사용됩니다. 이러한 설정에서 암호와 같은 민감한 정보를 제공하려면 AWS Secrets Manager에서 민감한 구성 데이터 저장을 참조하세요.

사용 가능한 구성 분류는 Amazon EMR 릴리스 버전에 따라 달라집니다. 특정 릴리스 버전에서 지원되는 구성 분류 목록은 Amazon EMR 릴리스에 대한 정보 아래 해당 릴리스 버전 페이지를 참조하세요.

다음은 구성 목록의 예제 JSON 파일입니다.

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.security.groups.cache.secs": "250"
    }
  },
  {
    "Classification": "mapred-site",
    "Properties": {
      "mapred.tasktracker.map.tasks.maximum": "2",
      "mapreduce.map.sort.spill.percent": "0.90",
      "mapreduce.tasktracker.reduce.tasks.maximum": "5"
    }
  }
]
```

구성 분류는 애플리케이션별 구성 파일에 매핑되는 경우가 많습니다. 예를 들어, `hive-site` 분류는 Hive용 `hive-site.xml` 구성 파일의 설정에 매핑됩니다. 이 작업에 대한 예외는, 더 이상 사용되지 않는 부트스트랩 작업 `configure-daemons`입니다. 이 작업은 환경 파라미터(예: `--namenode-heap-`

size)를 설정하는 데 사용됩니다. 이와 같은 옵션은 고유의 중첩 내보내기 분류를 갖춘 hadoop-env 및 yarn-env 분류에 포함됩니다. 분류가 env로 끝나는 경우 내보내기 하위 분류를 사용합니다.

또 다른 한 가지 예외는 s3get입니다. 이 항목은 클라이언트 측 암호화에 사용할 클러스터의 각 노드에 고객 EncryptionMaterialsProvider 객체를 배치하는 데 사용됩니다. 이 용도로 emrfs-site 분류에 옵션이 추가되었습니다.

다음은 hadoop-env 분류의 예제입니다.

```
[
  {
    "Classification": "hadoop-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "HADOOP_DATANODE_HEAPSIZE": "2048",
          "HADOOP_NAMENODE_OPTS": "-XX:GCTimeRatio=19"
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

다음은 yarn-env 분류의 예제입니다.

```
[
  {
    "Classification": "yarn-env",
    "Properties": {

    },
    "Configurations": [
      {
        "Classification": "export",
        "Properties": {
          "YARN_RESOURCEMANAGER_OPTS": "-Xdebug -Xrunjdwp:transport=dt_socket"
```

```
        },
        "Configurations": [

        ]
      }
    ]
  }
]
```

다음 설정은 구성 파일에 속하지 않지만 Amazon EMR에서 사용자를 대신하여 여러 설정을 구성하기 위해 사용됩니다.

Amazon EMR에서 큐레이트된 설정

Application	릴리스 레이블 분류	유효한 속성	사용해야 하는 경우
Spark	spark	maximizeResourceAllocation	각 노드의 리소스를 최대한 활용할 수 있도록 실행기를 구성합니다.

주제

- [클러스터 생성 시 애플리케이션 구성]
- [실행 중인 클러스터에서 인스턴스 그룹 재구성]
- [AWS Secrets Manager에 민감한 구성 데이터 저장]
- [특정 Java 가상 머신을 사용하도록 애플리케이션 구성]

클러스터 생성 시 애플리케이션 구성

클러스터를 생성할 때 Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여 애플리케이션의 기본 구성을 재정의할 수 있습니다.

애플리케이션의 기본 구성을 재정의하려면 구성 분류에 사용자 지정 값을 지정합니다. 구성 분류는 애플리케이션의 구성 XML 파일(예: `hive-site.xml`)에 해당합니다.

구성 분류는 Amazon EMR 릴리스 버전에 따라 다릅니다. 특정 릴리스 버전에서 사용할 수 있는 구성 분류 목록은 릴리스 세부 정보 페이지를 참조하세요. 예를 들어 [Amazon EMR 릴리스 6.4.0]과 같습니다.

클러스터를 생성할 때 콘솔에서 구성 제공

구성을 제공하려면 클러스터 생성 페이지로 이동하여 소프트웨어 설정을 확장합니다. 그런 다음 JSON 또는 콘솔에 섀도우 텍스트로 표시되는 간편 구문을 사용하여 구성을 직접 입력할 수 있습니다. 그렇지 않으면 파일의 Amazon S3 URI를 JSON `Configurations` 객체와 함께 제공할 수 있습니다.

인스턴스 그룹에 대한 구성을 제공하려면 클러스터 목록에서 클러스터를 선택한 다음, 구성 탭을 선택합니다. 인스턴스 그룹 구성 테이블에서 편집할 인스턴스 그룹을 선택한 다음, 재구성을 선택합니다.

클러스터를 생성할 AWS CLI 때를 사용하여 구성 제공

로컬로 또는 Amazon S3에 저장된 JSON 파일의 경로를 제공하여 create-cluster에 대한 구성을 제공할 수 있습니다. 다음 예제에서는 Amazon EMR 기본 역할을 사용하며, 이 역할이 생성되어 있다고 가정합니다. 하지만 역할을 직접 생성해야 하는 경우에는 먼저 `aws emr create-default-roles`를 실행하십시오.

구성이 로컬 디렉터리에 있는 경우 다음 예제 명령을 사용할 수 있습니다.

```
aws emr create-cluster --use-default-roles --release-label emr-7.9.0 --applications
 Name=Hive \
--instance-type m5.xlarge --instance-count 3 --configurations file://./
configurations.json
```

구성이 Amazon S3 경로에 있는 경우 Amazon S3 경로를 `create-cluster` 명령에 전달하기 전에 다음 해결 방법을 설정해야 합니다.

```
#!/bin/sh
# Assume the ConfigurationS3Path is not public, and its present in the same AWS account
 as the EMR cluster
ConfigurationS3Path="s3://amzn-s3-demo-bucket/config.json"
# Get a presigned HTTP URL for the s3Path
ConfigurationURL=`aws s3 presign $ConfigurationS3Path --expires-in 300`
# Fetch the presigned URL, and minify the JSON so that it spans only a single line
Configurations=`curl $ConfigurationURL | jq -c .`
aws emr create-cluster --use-default-roles --release-label emr-5.34.0 --instance-type
 m5.xlarge --instance-count 2 --applications Name=Hadoop Name=Spark --configurations
 $Configurations
```

클러스터를 생성할 때 Java SDK를 사용하여 구성 제공

다음 프로그램 발췌문에서는 AWS SDK for Java를 사용하여 구성을 제공하는 방법을 보여줍니다.

```
Application hive = new Application().withName("Hive");

Map<String,String> hiveProperties = new HashMap<String,String>();
 hiveProperties.put("hive.join.emit.interval","1000");
 hiveProperties.put("hive.merge.mapfiles","true");

Configuration myHiveConfig = new Configuration()
 .withClassification("hive-site")
 .withProperties(hiveProperties);

RunJobFlowRequest request = new RunJobFlowRequest()
 .withName("Create cluster with ReleaseLabel")
 .withReleaseLabel("emr-5.20.0")
 .withApplications(hive)
 .withConfigurations(myHiveConfig)
 .withServiceRole("EMR_DefaultRole")
 .withJobFlowRole("EMR_EC2_DefaultRole")
 .withInstances(new JobFlowInstancesConfig()
  .withEc2KeyName("myEc2Key")
  .withInstanceCount(3)
  .withKeepJobFlowAliveWhenNoSteps(true)
  .withMasterInstanceType("m4.large")
  .withSlaveInstanceType("m4.large")
 );
```

실행 중인 클러스터에서 인스턴스 그룹 재구성

Amazon EMR 버전 5.21.0 이상에서는 클러스터 애플리케이션을 재구성하고 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. 이를 위해 Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 SDK를 AWS 사용할 수 있습니다.

새 Amazon EMR 콘솔에서 인스턴스 그룹의 애플리케이션 구성을 업데이트하면 콘솔은 새 구성을 기존 구성과 병합하여 새로운 활성 구성을 생성하려고 시도합니다. Amazon EMR에서 구성을 병합할 수 없는 특이한 경우 콘솔에서 경고를 표시합니다.

인스턴스 그룹의 재구성 요청을 제출한 후 Amazon EMR에서는 새 구성 사양에 버전 번호를 할당합니다. CloudWatch 이벤트를 조회하여 구성의 버전 번호나 인스턴스 그룹의 상태를 추적할 수 있습니다. 자세한 내용은 CloudWatch 이벤트 모니터링을 참조하십시오.

> **Note**
>
> 클러스터 생성 중에 지정된 클러스터 구성만 재정의할 수 있고 삭제할 수는 없습니다. 기존 구성과 제공한 파일 간에 차이가 있는 경우 Amazon EMR은 수동으로 수정한 구성(예: SSH를 사용하여 클러스터에 연결하는 동안 수정한 구성)을 지정된 인스턴스 그룹의 클러스터 기본값으로 재설정합니다.

인스턴스 그룹을 재구성할 때 고려 사항

재구성 작업

Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여 재구성 요청을 제출하면 Amazon EMR은 기존 클러스터 내 구성 파일을 확인합니다. 기존 구성과 제공한 파일 간에 차이가 있는 경우 Amazon EMR은 재구성 작업을 시작하고 일부 애플리케이션을 다시 시작하며 수동으로 수정한 구성(예: SSH를 사용하여 클러스터에 연결하는 동안 수정한 구성)을 지정된 인스턴스 그룹의 클러스터 기본값으로 재설정합니다.

> **Note**
>
> Amazon EMR은 모든 인스턴스 그룹 재구성 중에 몇 가지 기본 작업을 수행합니다. 이러한 기본 작업은 이전에 수행한 클러스터 사용자 지정과 충돌하여 재구성 실패로 이어질 수 있습니다. 재구성 실패 문제 해결 방법에 대한 자세한 내용은 인스턴스 그룹 재구성 문제 해결 섹션을 참조하세요.

또한 Amazon EMR은 요청에서 지정한 구성 분류에 대한 재구성 작업을 시작합니다. 이러한 작업의 전체 목록은 사용 중인 Amazon EMR 버전의 구성 분류 섹션을 참조하세요. 예를 들어, 6.2.0 구성 분류와 같습니다.

> **Note**
>
> Amazon EMR 릴리스 안내서에는 Amazon EMR 버전 5.32.0 및 6.2.0부터 시작하는 재구성 작업만 나와 있습니다.

서비스 중단

Amazon EMR은 롤링 프로세스에 따라 태스크 및 코어 인스턴스 그룹의 인스턴스를 재구성합니다. 인스턴스 그룹에 있는 인스턴스 중 10%만 동시에 수정 및 다시 시작됩니다. 이 프로세스는 끝내는 데 오래 걸리지만 실행 중인 클러스터에서 애플리케이션이 실패할 가능성을 줄여 줍니다.

YARN을 다시 시작하는 중에 YARN 작업을 실행하려면 여러 프라이머리 노드가 있는 Amazon EMR 클러스터를 생성하거나 `yarn-site` 구성 분류에서 `yarn.resourcemanager.recovery.enabled`를 `true`로 설정하면 됩니다. 여러 프라이머리 노드 사용에 대한 자세한 내용은 [고가용성 YARN ResourceManager]를 참조하세요.

애플리케이션 검증

Amazon EMR은 재구성 다시 시작 프로세스 후에 클러스터에서 각 애플리케이션이 실행 중인지 확인합니다. 애플리케이션을 사용할 수 없는 경우 전체 재구성 작업이 실패합니다. 재구성이 실패하면 Amazon EMR은 구성 파라미터를 이전에 작동하던 버전으로 되돌립니다.

> **Note**
>
> 재구성 실패를 방지하려면 사용하려는 클러스터에만 애플리케이션을 설치하는 것이 좋습니다. 또한 재구성 요청을 제출하기 전에 모든 클러스터 애플리케이션이 정상이고 실행 중인지 확인하는 것이 좋습니다.

재구성 유형

다음 두 가지 방법으로 인스턴스 그룹을 재구성할 수 있습니다.

- 덮어쓰기. 기본 재구성 방법이며 5.35.0 및 6.6.0 이전의 Amazon EMR 릴리스에서 사용할 수 있는 유일한 방법입니다. 이 재구성 방법은 클러스터 내 파일을 새로 제출한 구성 세트로 무차별적으로 덮어씁니다. 이 메서드는 재구성 API 외부에서 수행된 구성 파일 변경 내용을 모두 지웁니다.
- 병합. Amazon EMR 릴리스 5.35.0 및 6.6.0 이상에서 지원되는 재구성 방법입니다. 단, Amazon EMR 콘솔에서는 이를 지원하는 버전이 없으므로 예외입니다. 이 재구성 방법은 새로 제출된 구성을 클러스터에 이미 있는 구성과 병합합니다. 이 옵션은 제출한 새 구성을 추가하거나 수정하기만 합니다. 기존 구성을 보존합니다.

> **Note**
>
> Amazon EMR은 서비스가 올바르게 실행되도록 하는 데 필요한 일부 필수 Hadoop 구성을 계속 덮어씁니다.

제한 사항

실행 중인 클러스터에서 인스턴스 그룹을 재구성할 경우 다음 제한 사항을 고려합니다.

- Yarn 기반이 아닌 애플리케이션에서 특히 애플리케이션이 제대로 구성되지 않은 경우 다시 시작 중에 실패하거나 클러스터 문제가 발생할 수 있습니다. 클러스터의 최대 메모리 및 CPU 사용량에 근접하면 다시 시작 프로세스 후에 문제가 발생할 수 있습니다. 프라이머리 인스턴스 그룹의 경우 특히 그렇습니다.
- 인스턴스 그룹 크기가 조정되는 동안에는 재구성 요청을 제출할 수 없습니다. 인스턴스 그룹의 크기를 조정하는 동안 재구성이 시작되면 인스턴스 그룹의 크기 조정이 완료될 때까지 재구성을 시작할 수 없으며, 그 반대의 경우도 마찬가지입니다.
- 인스턴스 그룹을 재구성한 후 Amazon EMR에서 애플리케이션을 다시 시작하여 새 구성을 적용합니다. 재구성하는 동안 애플리케이션이 사용 중이면 작업이 실패하거나 예기치 않은 다른 애플리케이션 동작이 발생할 수 있습니다.
- 인스턴스 그룹의 재구성이 실패하면 Amazon EMR이 구성 파라미터를 이전에 작동하던 버전으로 되돌립니다. 리버전 프로세스가 실패하면 새 `ModifyInstanceGroup` 요청을 제출하여 인스턴스 그룹을 `SUSPENDED` 상태에서 복구해야 합니다.
- Phoenix 구성 분류에 대한 재구성 요청은 Amazon EMR 버전 5.23.0 이상에서만 지원되고 Amazon EMR 버전 5.21.0 또는 5.22.0에서는 지원되지 않습니다.
- HBase 구성 분류에 대한 재구성 요청은 Amazon EMR 버전 5.30.0 이상에서만 지원되고 Amazon EMR 버전 5.23.0~5.29.0에서는 지원되지 않습니다.
- Amazon EMR은 Amazon EMR 버전 5.27.0 이상에서만 여러 프라이머리 노드가 있는 Amazon EMR 클러스터에 대한 애플리케이션 재구성 요청을 지원합니다.
- 모든 Hadoop KMS 구성 분류 또는 `hdfs-encryption-zones` 분류의 재구성은 여러 프라이머리 노드가 있는 Amazon EMR 클러스터에서 지원되지 않습니다.
- Amazon EMR은 YARN ResourceManager를 다시 시작해야 하는 용량 스케줄러에 대한 특정 재구성 요청을 지원하지 않습니다. 예를 들어 대기열을 완전히 제거할 수는 없습니다.

콘솔에서 인스턴스 그룹 재구성

> **Note**
>
> Amazon EMR 콘솔은 병합 유형 재구성을 지원하지 않습니다.

1. https://console.aws.amazon.com/elasticmapreduce/에서 Amazon EMR 콘솔을 엽니다.
2. 클러스터 목록의 이름 아래에서 재구성할 활성 클러스터를 선택합니다.
3. 클러스터의 클러스터 세부 정보 페이지를 열고 구성 탭으로 이동합니다.
4. 필터 드롭다운 목록에서 재구성할 인스턴스 그룹을 선택합니다.
5. 재구성 드롭다운 메뉴에서 테이블에서 편집 또는 JSON 파일에서 편집을 선태합니다.

    - 테이블에서 편집 - 구성 분류 테이블에서 기존 구성의 속성과 값을 편집하거나 구성 추가를 선택하여 추가 구성 분류를 제공합니다.
    - JSON 파일에서 편집 - JSON에 직접 구성을 입력하거나 섀도우 텍스트로 표시된 간편 구문을 사용합니다. 그렇지 않으면 파일의 Amazon S3 URI를 JSON `Configurations` 객체와 함께 제공합니다.

    > **Note**
    >
    > 구성 분류 테이블의 소스 열은 클러스터를 생성할 때 또는 이 인스턴스 그룹의 추가 구성을 지정할 때 구성이 제공되는지 여부를 보여줍니다. 두 소스에서 모두 인스턴스 그룹의 구성을 편집할 수 있습니다. 초기 클러스터 구성을 삭제할 수 없지만 인스턴스 그룹의 초기 클러스터 구성을 재정의할 수는 있습니다.
    > 테이블에서 직접 중첩된 구성 분류를 추가하거나 편집할 수도 있습니다. 예를 들어, `hadoop-env`의 `export` 하위 분류를 추가로 제공하려면 표에서 `hadoop.export` 구성 분류를 추가합니다. 그런 다음 이 분류의 특정 속성과 값을 제공합니다.

6. (선택 사항) Apply this configuration to all active instance groups(모든 활성 인스턴스 그룹에 이 구성 적용)를 선택합니다.
7. 변경 사항을 저장합니다.

CLI를 사용하여 인스턴스 그룹 재구성

modify-instance-groups 명령을 사용하여 실행 중인 클러스터의 인스턴스 그룹에 대해 새 구성을 지정합니다.

> **Note**
>
> 다음 예제에서 *<j-2AL4XXXXXX5T9>*를 클러스터 ID로 바꾸고, *<ig-1xxxxxxx9>*를 인스턴스 그룹 ID로 바꿉니다.

Example - 인스턴스 그룹의 구성 대치

다음 예제에서는 구성 JSON 파일 `instanceGroups.json`을 참조하여 인스턴스 그룹에 대해 YARN NodeManager 디스크 상태 검사기의 속성을 편집합니다.

1. 구성 분류를 준비하고 명령을 실행할 디렉터리와 동일한 디렉터리에 `instanceGroups.json`으로 저장합니다.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. 다음 명령을 실행합니다.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
```

```
--instance-groups file://instanceGroups.json
```

Example - 인스턴스 그룹에 구성 추가

인스턴스 그룹에 구성을 추가하려면 새 `ModifyInstanceGroup` 요청에 해당 인스턴스 그룹에 대해 이전에 지정한 구성도 포함해야 합니다. 그렇지 않으면 이전에 지정한 구성이 제거됩니다.

다음 예제에서는 YARN NodeManager 가상 메모리 검사기의 속성을 추가합니다. 또한 구성에는 값을 덮어쓰지 않도록 YARN NodeManager 디스크 상태 검사기에 대해 이전에 지정한 값도 포함됩니다.

1. `instanceGroups.json`에서 다음 콘텐츠를 준비하고 명령을 실행할 동일한 디렉터리에 저장합니다.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0",
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. 다음 명령을 실행합니다.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example - 병합 유형 재구성을 사용하여 인스턴스 그룹에 구성 추가

기본 덮어쓰기 재구성 방법을 사용하여 구성을 추가하려면 해당 인스턴스 그룹에 대해 이전에 지정한 모든 구성을 새 `ModifyInstanceGroup` 요청에 포함해야 합니다. 그렇지 않으면 덮어쓰기로 이전에 지정한 구성이 제거됩니다. 병합 재구성에서는 이 작업을 수행할 필요가 없습니다. 대신 요청에 새 구성만 포함되도록 해야 합니다.

다음 예제에서는 YARN NodeManager 가상 메모리 검사기의 속성을 추가합니다. 이는 병합 유형 재구성이므로 YARN NodeManager 디스크 상태 검사기에 대해 이전에 지정된 값을 덮어쓰지 않습니다.

1. `instanceGroups.json`에서 다음 콘텐츠를 준비하고 명령을 실행할 동일한 디렉터리에 저장합니다.

```
[
   {"InstanceGroupId":"<ig-1xxxxxxx9>",
    "ReconfigurationType" :"MERGE",
      "Configurations":[
         {"Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. 다음 명령을 실행합니다.

```
aws emr modify-instance-groups --cluster-id <j-2AL4XXXXXX5T9> \
--instance-groups file://instanceGroups.json
```

Example - 인스턴스 그룹의 구성 삭제

인스턴스 그룹에 대한 구성을 삭제하려면 이전 구성을 제외하는 새 재구성 요청을 제출합니다.

> **Note**
>
> 초기 클러스터 구성은 재정의만 가능합니다. 삭제할 수는 없습니다.

예를 들어, 이전 예제에서 YARN NodeManager 디스크 상태 검사기의 구성을 삭제하려면 다음 콘텐츠를 포함하는 새 `instanceGroups.json`을 제출합니다.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Properties":{
               "yarn.nodemanager.vmem-check-enabled":"true",
               "yarn.nodemanager.vmem-pmem-ratio":"3.0"
            },
            "Configurations":[]
         }
      ]
   }
]
```

> **Note**
>
> 마지막 재구성 요청에서 모든 구성을 삭제하려면 빈 구성 배열이 있는 재구성 요청을 제출합니다. 예:
>
> ```
> [
>    {
>       "InstanceGroupId":"<ig-1xxxxxxx9>",
>       "Configurations":[]
>    }
> ]
> ```

Example - 하나의 요청으로 인스턴스 그룹 재구성 및 크기 조정

다음 예제 JSON에서는 동일한 요청의 인스턴스 그룹을 재구성하고 그 크기를 조정하는 방법을 보여줍니다.

```
[
   {
      "InstanceGroupId":"<ig-1xxxxxxx9>",
      "InstanceCount":5,
      "EC2InstanceIdsToTerminate":["i-123"],
      "ForceShutdown":true,
      "ShrinkPolicy":{
         "DecommissionTimeout":10,
         "InstanceResizePolicy":{
            "InstancesToTerminate":["i-123"],
            "InstancesToProtect":["i-345"],
            "InstanceTerminationTimeout":20
         }
      },
      "Configurations":[
         {
            "Classification":"yarn-site",
            "Configurations":[],
            "Properties":{
               "yarn.nodemanager.disk-health-checker.enable":"true",
               "yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-
percentage":"100.0"
            }
         }
      ]
   }
]
```

Java SDK를 사용하여 인스턴스 그룹 재구성

Note

다음 예제에서 *<j-2AL4XXXXXX5T9>*를 클러스터 ID로 바꾸고, *<ig-1xxxxxxx9>*를 인스턴스 그룹 ID로 바꿉니다.

다음 코드 스니펫은 AWS SDK for Java를 사용하는 인스턴스 그룹의 새 구성을 제공합니다.

```
AWSCredentials credentials = new BasicAWSCredentials("access-key", "secret-key");
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);

Map<String,String> hiveProperties = new HashMap<String,String>();
hiveProperties.put("hive.join.emit.interval","1000");
hiveProperties.put("hive.merge.mapfiles","true");

Configuration configuration = new Configuration()
    .withClassification("hive-site")
    .withProperties(hiveProperties);

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withReconfigurationType("MERGE");
    .withConfigurations(configuration);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

다음 코드 스니펫은 빈 구성 배열을 제공하여 이전에 지정한 인스턴스 그룹 구성을 삭제합니다.

```
List<Configuration> configurations = new ArrayList<Configuration>();

InstanceGroupModifyConfig igConfig = new InstanceGroupModifyConfig()
    .withInstanceGroupId("<ig-1xxxxxxx9>")
    .withConfigurations(configurations);

ModifyInstanceGroupsRequest migRequest = new ModifyInstanceGroupsRequest()
    .withClusterId("<j-2AL4XXXXXX5T9>")
    .withInstanceGroups(igConfig);

emr.modifyInstanceGroups(migRequest);
```

인스턴스 그룹 재구성 문제 해결

인스턴스 그룹의 재구성 프로세스가 실패할 경우 Amazon EMR은 Amazon CloudWatch 이벤트를 사용하여 재구성을 되돌리고 실패 메시지를 로깅합니다. 이벤트는 재구성 실패에 대한 간략한 요약을 제공합니다. 재구성이 실패한 인스턴스와 해당 실패 메시지를 나열합니다. 다음은 실패 메시지의 예입니다.

```
The reconfiguration operation for instance group ig-1xxxxxxx9 in Amazon EMR
 cluster j-2AL4XXXXXX5T9 (ExampleClusterName)
failed at 2021-01-01 00:00 UTC and took 2 minutes to fail. Failed configuration version
 is example12345.
Failure message: Instance i-xxxxxxx1, i-xxxxxxx2, i-xxxxxxx3 failed with message "This
 is an example failure message".
```

재구성 실패에 대한 추가 데이터를 수집하려면 노드 프로비저닝 로그를 확인할 수 있습니다. 이렇게 하면 다음과 같은 메시지를 수신할 때 특히 유용합니다.

```
i-xxxxxxx1 failed with message “Unable to complete transaction and some changes were
 applied.”
```

On the node

노드에 연결하여 노드 프로비저닝 로그에 액세스하는 방법

1. SSH를 사용하여 재구성이 실패한 노드에 연결합니다. 관련 지침은 Amazon EC2 Linux 인스턴스용 사용 설명서에서 거 Linux 인스턴스에 연결 을 참조하세요.
2. 노드 프로비저닝 로그 파일이 있는 다음 디렉터리로 이동합니다.

   ```
   /mnt/var/log/provision-node/
   ```

3. `reports` 하위 디렉터리를 열고 재구성을 위한 노드 프로비저닝 보고서를 검색합니다. `reports` 디렉터리는 재구성 버전 번호, 범용 고유 식별자(UUID), Amazon EC2 인스턴스 IP 주소 및 타임스탬프별로 로그를 구성합니다. 각 보고서는 재구성 프로세스에 대한 자세한 정보를 포함하는 압축된 YAML 파일입니다.

   다음은 보고서 파일 이름 및 경로에 대한 예제입니다.

   ```
   /reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
   xxx.ec2.internal/202104061715.yaml.gz
   ```

4. 다음 예제와 같이 `zless` 등의 파일 뷰어를 사용하여 보고서를 검토할 수 있습니다.

```
zless 202104061715.yaml.gz
```

Amazon S3

Amazon S3를 사용하여 노드 프로비저닝 로그에 액세스하는 방법

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/s3/ Amazon S3 콘솔을 엽니다.
2. 로그 파일을 아카이브하도록 클러스터를 구성할 때 지정한 Amazon S3 버킷을 엽니다.
3. 노드 프로비저닝 로그 파일이 있는 다음 폴더로 이동합니다.

```
amzn-s3-demo-bucket/elasticmapreduce/<cluster id>/node/<instance id>/provision-node/
```

4. `reports` 폴더를 열고 재구성을 위한 노드 프로비저닝 보고서를 검색합니다. `reports` 폴더는 재구성 버전 번호, 범용 고유 식별자(UUID), Amazon EC2 인스턴스 IP 주소 및 타임스탬프별로 로그를 구성합니다. 각 보고서는 재구성 프로세스에 대한 자세한 정보를 포함하는 압축된 YAML 파일입니다.

   다음은 보고서 파일 이름 및 경로에 대한 예제입니다.

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-xxx.ec2.internal/202104061715.yaml.gz
```

5. 로그 파일을 보려면 Amazon S3의에서 로컬 컴퓨터로 로그 파일을 텍스트 파일로 다운로드하면 됩니다. 관련 지침은 객체 다운로드를 참조하세요.

각 로그 파일에는 관련 재구성에 대한 자세한 프로비저닝 보고서가 들어 있습니다. 오류 메시지 정보를 찾으려면 보고서의 `err` 로그 수준을 검색하면 됩니다. 보고서 형식은 클러스터의 Amazon EMR 버전에 따라 다릅니다.

다음 예제에서는 5.32.0 및 6.2.0 이전의 Amazon EMR 릴리스 버전에 대한 오류 정보를 보여줍니다.

```
- !ruby/object:Puppet::Util::Log
      level: !ruby/sym err
      tags:
        - err
```

```
    message: "Example detailed error message."
    source: Puppet
    time: 2021-01-01 00:00:00.000000 +00:00
```

Amazon EMR 릴리스 버전 5.32.0 및 6.2.0 이상에서는 다음 형식을 대신 사용합니다.

```
- level: err
  message: 'Example detailed error message.'
  source: Puppet
  tags:
  - err
  time: '2021-01-01 00:00:00.000000 +00:00'
  file:
  line:
```

AWS Secrets Manager에 민감한 구성 데이터 저장

Amazon EMR은 사용자 지정 구성 데이터(예: `DescribeCluster` 및 `ListInstanceGroups`)를 생성하는 API 작업을 일반 텍스트로 설명하고 나열합니다. Amazon EMR은 AWS Secrets Manager 와 통합되어 Secrets Manager에 데이터를 저장하고 구성에서 보안 암호 ARN을 사용할 수 있습니다. 이렇게 하면 민감한 구성 데이터를 일반 텍스트로 Amazon EMR에 전달하여 외부 API에 노출하지 않습니다. 키-값 페어에 Secrets Manager에 저장된 보안 암호의 ARN이 포함되어 있다고 지정한 경우 Amazon EMR은 구성 데이터를 클러스터로 전송할 때 이 보안 암호를 검색합니다. Amazon EMR은 외부 API를 사용하여 구성을 표시할 때 주석을 전송하지 않습니다.

보안 암호 생성

보안 암호를 생성하려면 AWS Secrets Manager 사용 설명서에서 AWS Secrets Manager 보안 암호 생성의 단계를 따릅니다. 3단계에서는 일반 텍스트 필드를 선택하여 민감한 값을 입력해야 합니다.

Secrets Manager에서는 보안 암호에 최대 65,536바이트를 포함할 수 있지만 Amazon EMR은 속성 키(주석 제외) 및 검색된 보안 암호 값의 총 길이를 1,024자로 제한합니다.

Amazon EMR에 보안 암호를 검색할 액세스 권한 부여

Amazon EMR은 IAM 서비스 역할을 사용하여 클러스터를 프로비저닝하고 관리합니다. Amazon EMR의 서비스 역할은 리소스를 프로비저닝하고 클러스터 내에서 실행 중인 Amazon EC2 인스턴스 컨텍스트에서 수행되지 않는 서비스 수준 작업을 수행하는 경우 허용되는 작업을 정의합니다. 서비스 역할에 대한 자세한 내용은 Amazon EMR의 서비스 역할(EMR 역할) 및 IAM 역할 사용자 지정을 참조하세요.

Amazon EMR이 Secrets Manager로부터 보안 암호 값을 검색할 수 있도록 하려면 클러스터를 시작할 때 Amazon EMR 역할에 다음 정책 명령문을 추가합니다.

```
{
   "Sid":"AllowSecretsRetrieval",
   "Effect":"Allow",
   "Action":"secretsmanager:GetSecretValue",
   "Resource":[
      "arn:aws:secretsmanager:<region>:<aws-account-id>:secret:<secret-name>"
   ]
}
```

고객 관리형 로 보안 암호를 생성하는 경우 사용하는 키에 대한 Amazon EMR 역할에 `kms:Decrypt` 권한도 추가해야 AWS KMS key합니다. 자세한 내용은 AWS Secrets Manager User Guide의 [Authentication and access control for AWS Secrets Manager](#)를 참조하세요.

구성 분류에서 보안 암호 사용

구성 속성에 `EMR.secret@` 주석을 추가하여 해당 키-값 페어에 Secrets Manager에 저장된 보안 암호의 ARN이 포함되어 있음을 나타낼 수 있습니다.

다음 예제에서는 구성 분류에서 보안 암호 ARN을 제공하는 방법을 보여줍니다.

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
      "EMR.secret@presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
   }
}
```

클러스터를 생성하고 주석이 달린 구성을 제출하면 Amazon EMR이 구성 속성을 검증합니다. 구성이 유효하면 Amazon EMR은 구성에서 주석을 제거하고 Secrets Manager에서 보안 암호를 검색하여 클러스터에 적용하기 전에 실제 구성을 생성합니다.

```
{
   "Classification":"core-site",
   "Properties":{
      "presto.s3.access-key":"<sensitive-access-key>",
```

```
        "presto.s3.secret-key":"<my-secret-key-retrieved-from-Secrets-Manager>"
    }
}
```

DescribeCluster와 같은 작업을 직접 호출하면 Amazon EMR은 클러스터에서 현재 애플리케이션 구성을 반환합니다. 애플리케이션 구성 속성이 보안 암호 ARN을 포함하는 것으로 표시된 경우 DescribeCluster 직접 호출에서 반환된 애플리케이션 구성에는 보안 암호 값이 아닌 ARN이 포함됩니다. 이렇게 하면 클러스터에서만 보안 암호 값을 볼 수 있습니다.

```
{
    "Classification":"core-site",
    "Properties":{
        "presto.s3.access-key":"<sensitive-access-key>",
        "presto.s3.secret-key":"arn:aws:secretsmanager:<region>:<aws-account-
id>:secret:<secret-name>"
    }
}
```

보안 암호 값 업데이트

Amazon EMR은 연결된 인스턴스 그룹이 시작 중, 재구성 중 또는 조정 중일 때마다 주석이 달린 구성에서 보안 암호 값을 검색합니다. Secrets Manager를 사용하여 실행 중인 클러스터의 구성에 사용되는 보안 암호 값을 수정할 수 있습니다. 이렇게 하면 업데이트된 값을 수신할 각 인스턴스 그룹에 재구성 요청을 제출할 수 있습니다. 인스턴스 그룹을 재구성하는 방법 및 재구성 시 고려할 사항에 대한 자세한 내용은 실행 중인 클러스터에서 인스턴스 그룹 재구성 섹션을 참조하세요.

특정 Java 가상 머신을 사용하도록 애플리케이션 구성

Amazon EMR 릴리스에는 다양한 기본 Java 가상 머신(JVM) 버전이 있습니다. 이 페이지에서는 다양한 릴리스 및 애플리케이션에 대한 JVM 지원을 설명합니다.

고려 사항

애플리케이션에 지원되는 Java 버전에 대한 자세한 내용은 Amazon EMR 릴리스 안내서의 애플리케이션 페이지를 참조하세요.

- Amazon EMR은 클러스터에서 하나의 런타임 버전만 실행할 수 있으며, 동일한 클러스터의 다른 런타임 버전에서 다른 노드 또는 애플리케이션 실행을 지원하지 않습니다.

- Amazon EMR 7.x의 경우 Apache Livy를 제외하고 Java 17을 지원하는 애플리케이션에 대한 기본 JVM(Java Virtual Machine)은 Java 17입니다. 애플리케이션에 지원되는 JDK 버전에 대한 자세한 내용은 Amazon EMR 릴리스 안내서의 해당 릴리스 페이지를 참조하세요.
- Amazon EMR 7.1.0부터 Flink에서 지원하며 기본적으로 Java 17로 설정됩니다. 다른 버전의 Java 런타임을 사용하려면 `flink-conf`의 설정을 재정의합니다. Java 8 또는 Java 11을 사용하도록 Flink를 구성하는 방법에 대한 자세한 내용은 Java 11과 함께 실행하도록 Flink 구성을 참조하세요.
- Amazon EMR 5.x 및 6.x의 경우 기본 JVM(Java Virtual Machine)은 Java 8입니다.
  - Amazon EMR 릴리스 6.12.0 이상의 경우 일부 애플리케이션은 Java 11 및 17도 지원합니다.
  - Amazon EMR 릴리스 6.9.0 이상의 경우 Trino는 Java 17을 기본으로 지원합니다. Trino를 사용하는 Java 17에 대한 자세한 내용은 Trino 블로그에서 Trino updates to Java 17을 참조하세요.

런타임 버전을 선택할 때는 다음과 같은 애플리케이션별 고려 사항에 주의합니다.

애플리케이션별 Java 구성 참고 사항

Application	Java 구성 참고 사항
Spark	기본이 아닌 Java 버전으로 Spark를 실행하려면 Spark 및 Hadoop을 모두 구성해야 합니다. 예시는 JVM 재정의 섹션을 참조하세요. • 기본 인스턴스 프로세스의 Java 런타임을 업데이트하도록 `spark-env` 에서 `JAVA_HOME` 을 구성합니다. 예를 들어, spark-submit, spark-shell, Spark 기록 서버가 이에 해당합니다. • Hadoop 구성을 수정하여 Spark 실행기와 YARN `ApplicationMaster` 의 Java 런타임을 업데이트합니다.
Spark RAPIDS	구성된 Spark용 Java 버전으로 RAPIDS를 실행할 수 있습니다.
Iceberg	Iceberg를 사용하는 애플리케이션의 구성된 Java 버전으로 Iceberg를 실행할 수 있습니다.

Application	Java 구성 참고 사항
델타	Delta를 사용하는 애플리케이션의 구성된 Java 버전으로 Delta를 실행할 수 있습니다.
Hudi	Hudi를 사용하는 애플리케이션의 구성된 Java 버전으로 Hudi를 실행할 수 있습니다.
Hadoop	Hadoop용 JVM을 업데이트하려면 `hadoop-env` 를 수정합니다. 예시는 JVM 재정의 섹션을 참조하세요.
Hive	Hive용 Java 버전을 11 또는 17로 설정하려면 Hadoop JVM 설정을 사용하려는 Java 버전으로 구성합니다.
HBase	HBase용 JVM을 업데이트하려면 `hbase-env` 를 수정합니다. 기본적으로 Amazon EMR은 `hbase-env` 에서 설정을 재정의하지 않는 한 Hadoop용 JVM 구성을 기반으로 HBase JVM을 설정합니다. 예시는 JVM 재정의 섹션을 참조하세요.
Flink	Flink용 JVM을 업데이트하려면 `flink-conf` 를 수정합니다. 기본적으로 Amazon EMR은 `flink-conf` 에서 설정을 재정의하지 않는 한 Hadoop용 JVM 구성을 기반으로 Flink JVM을 설정합니다. 자세한 내용은 Java 11과 함께 실행하도록 Flink 구성 단원을 참조하십시오.
Oozie	Java 11 또는 17에서 실행하도록 Oozie를 구성하려면 Oozie 서버, Oozie LauncherAM Launcher AM을 구성하고 클라이언트 측 실행 파일 및 작업 구성을 변경합니다. Java 17에서 실행하도록 `EmbeddedOozieServer` 를 구성할 수도 있습니다. 자세한 내용은 Oozie용 자바 버전 구성 단원을 참조하십시오.

Application	Java 구성 참고 사항
Pig	Pig는 Java 8만 지원합니다. Java 11 또는 17을 Hadoop과 함께 사용하고 동일한 클러스터에서 Pig를 실행할 수 없습니다.

JVM 재정의

Amazon EMR 릴리스의 JVM 설정을 재정의하려면(예: Amazon EMR 릴리스 6.12.0을 사용하는 클러스터에서 Java 17을 사용하는 경우) 해당 `JAVA_HOME` 설정을 환경 분류(Flink를 제외한 모든 애플리케이션에서 *`application`*`-env`)에 제공합니다. Flink의 경우 환경 분류는 `flink-conf`입니다. Flink를 사용하여 Java 런타임을 구성하는 단계는 Java 11과 함께 실행하도록 Flink 구성 섹션을 참조하세요.

주제

- Apache Spark로 JVM 설정 재정의
- Apache HBase로 JVM 설정 재정의
- Apache Hadoop 및 Hive로 JVM 설정 재정의

Apache Spark로 JVM 설정 재정의

Amazon EMR 릴리스 6.12 이상에서 Spark를 사용하는 경우 실행기가 Java 11 또는 17을 사용하도록 환경을 설정할 수 있습니다. 또한 5.x 미만의 Amazon EMR 릴리스에서 Spark를 사용하고 클러스터 모드에서 제출용 드라이버를 작성하는 경우 드라이버는 Java 7을 사용합니다. 하지만 실행기에서 Java 8을 사용하도록 환경을 설정할 수 있습니다.

Spark용 JVM을 재정의하려면 하둡 및 Spark 분류를 모두 설정해야 합니다.

```
{
"Classification": "hadoop-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
            }
        ],
```

```
        "Properties": {}
    },
    {
"Classification": "spark-env",
        "Configurations": [
            {
"Classification": "export",
                "Configurations": [],
                "Properties": {
"JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
                }
            }
        ],
        "Properties": {}
    }
```

다음 예제에서는 모든 구성 요소에서 일관된 Java 버전 사용을 보장하기 위해 EMR 7.0.0+에 필요한 구성 파라미터를 추가하는 방법을 보여줍니다.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.executorEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0",
      "spark.yarn.appMasterEnv.JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
    }
  },
  {
    "Classification": "hadoop-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
  },
  {
    "Classification": "spark-env",
    "Configurations": [
```

```
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/java-1.8.0"
        }
      }
    ],
    "Properties": {}
  }
]
```

Apache HBase로 JVM 설정 재정의

Java 11을 사용하도록 HBase를 구성하려면 클러스터를 시작할 때 다음 구성을 설정하면 됩니다.

```
[
  {
    "Classification": "hbase-env",
    "Configurations": [
      {
        "Classification": "export",
        "Configurations": [],
        "Properties": {
          "JAVA_HOME": "/usr/lib/jvm/jre-11"
        }
      }
    ],
    "Properties": {}
  }
]
```

Apache Hadoop 및 Hive로 JVM 설정 재정의

다음 예제에서는 Hadoop 및 Hive용 JVM을 버전 17로 설정하는 방법을 보여줍니다.

```
[
    {
        "Classification": "hadoop-env",
            "Configurations": [
                {
```

```
                        "Classification": "export",
                        "Configurations": [],
                        "Properties": {
                            "JAVA_HOME": "/usr/lib/jvm/jre-17"
                        }
                    }
            ],
            "Properties": {}
    }
]
```

서비스 포트

다음은 YARN 및 HDFS 서비스 포트입니다. 이러한 설정은 하둡 기본값을 반영합니다. 다른 방식으로 문서화되지 않은 한, 기타 애플리케이션 서비스는 기본 포트에서 호스팅됩니다. 자세한 내용은 애플리케이션의 프로젝트 설명서를 참조하십시오.

YARN 및 HDFS에 대한 포트 설정

설정	호스트 이름/포트
`fs.default.name`	기본값(`hdfs://`*`emrDeterminedIP`* `:8020`)
`dfs.datanode.address`	기본값(`0.0.0.0:50010`)
`dfs.datanode.http.address`	기본값(`0.0.0.0:50075`)
`dfs.datanode.https.address`	기본값(`0.0.0.0:50475`)
`dfs.datanode.ipc.address`	기본값(`0.0.0.0:50020`)
`dfs.http.address`	기본값(`0.0.0.0:50070`)
`dfs.https.address`	기본값(`0.0.0.0:50470`)
`dfs.secondary.http.address`	기본값(`0.0.0.0:50090`)
`yarn.nodemanager.address`	기본값(`${yarn.nodemanager.hostname}:0`)
`yarn.nodemanager.localizer.address`	기본값(`${yarn.nodemanager.hostname}:8040`)

설정	호스트 이름/포트
yarn.nodemanager.webapp.address	기본값(${yarn.nodemanager.hostname }:8042)
yarn.resourcemanager.address	기본값(${yarn.resourcemanager.host name}:8032)
yarn.resourcemanager.admin.address	기본값(${yarn.resourcemanager.host name}:8033)
yarn.resourcemanager.resource-tracker.address	기본값(${yarn.resourcemanager.host name}:8031)
yarn.resourcemanager.scheduler.address	기본값(${yarn.resourcemanager.host name}:8030)
yarn.resourcemanager.webapp.address	기본값(${yarn.resourcemanager.host name}:8088)
yarn.web-proxy.address	기본값(no-value)
yarn.resourcemanager.hostname	*emrDeterminedIP*

Note

*emrDeterminedIP*는 Amazon EMR 컨트롤 플레인에서 생성된 IP 주소입니다. 최신 버전에서는 yarn.resourcemanager.hostname 및 fs.default.name 설정을 제외하고 이 규칙이 제거되었습니다.

애플리케이션 사용자

애플리케이션은 프로세스를 고유의 사용자로 실행합니다. 예를 들어, Hive JVM은 사용자 hive로 실행되고 MapReduce JVM은 mapred로 실행됩니다. 다음의 프로세스 상태 예제에서 이 과정을 볼 수 있습니다.

```
USER       PID %CPU %MEM    VSZ   RSS TTY      STAT START   TIME COMMAND
```

```
hive      6452  0.2  0.7 853684 218520 ?       Sl   16:32   0:13 /usr/lib/jvm/
java-openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-
metastore.log -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop
hive      6557  0.2  0.6 849508 202396 ?       Sl   16:32   0:09 /usr/lib/jvm/java-
openjdk/bin/java -Xmx256m -Dhive.log.dir=/var/log/hive -Dhive.log.file=hive-server2.log
 -Dhive.log.threshold=INFO -Dhadoop.log.dir=/usr/lib/hadoop/l
hbase     6716  0.1  1.0 1755516 336600 ?      Sl   Jun21   2:20 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_master -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hbase     6871  0.0  0.7 1672196 237648 ?      Sl   Jun21   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_thrift -XX:OnOutOfMemoryError=kill -9 %p -Xmx1024m -ea -XX:
+UseConcMarkSweepGC -XX:+CMSIncrementalMode -Dhbase.log.dir=/var/
hdfs      7491  0.4  1.0 1719476 309820 ?      Sl   16:32   0:22 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_namenode -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-hdfs -
Dhadoop.log.file=hadoop-hdfs-namenode-ip-10-71-203-213.log -Dhadoo
yarn      8524  0.1  0.6 1626164 211300 ?      Sl   16:33   0:05 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_proxyserver -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn -
Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-yarn-
yarn      8646  1.0  1.2 1876916 385308 ?      Sl   16:33   0:46 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_resourcemanager -Xmx1000m -Dhadoop.log.dir=/var/log/hadoop-yarn
 -Dyarn.log.dir=/var/log/hadoop-yarn -Dhadoop.log.file=yarn-y
mapred    9265  0.2  0.8 1666628 260484 ?      Sl   16:33   0:12 /usr/lib/jvm/java-
openjdk/bin/java -Dproc_historyserver -Xmx1000m -Dhadoop.log.dir=/usr/lib/hadoop/logs -
Dhadoop.log.file=hadoop.log -Dhadoop.home.dir=/usr/lib/hadoop
```

Amazon EMR 아티팩트 리포지토리를 사용하여 종속성 확인

Amazon EMR 아티팩트 리포지토리를 사용하여 Amazon EMR 릴리스 버전 5.18.0부터 특정 Amazon EMR 릴리스 버전에 사용할 수 있는 정확한 버전의 라이브러리와 종속성에 대해 Apache Hive 및 Apache Hadoop 작업 코드를 빌드할 수 있습니다. 리포지토리에 있는 Amazon EMR 아티팩트에 대해 빌드하면 작업이 빌드된 라이브러리의 버전이 클러스터에서 런타임에 제공되는 버전과 똑같아지므로 런타임 클래스 경로 문제를 피할 수 있습니다. 현재로서는 Maven 빌드에만 Amazon EMR 아티팩트를 사용할 수 있습니다.

아티팩트 리포지토리에 액세스하려면 리포지토리 URL을 Maven 설정 파일이나 특정 프로젝트의 `pom.xml` 구성 파일에 추가하십시오. 그런 다음 프로젝트 구성에서 종속성을 지정할 수 있습니다. 종속성 버전의 경우 Amazon EMR 5.x 릴리스 버전에서 원하는 릴리스의 구성 요소 버전 아래 나열된 버전을 사용합니다. 예를 들어 the section called “5.36.2 구성 요소 버전”에서 최신 Amazon EMR 릴리스의 구성 요소 버전을 볼 수 있습니다. 프로젝트의 아티팩트가 구성 요소 버전에 나열되어 있지 않으면 해당 릴리스의 Hive 및 Hadoop에 대해 나열된 버전을 지정하십시오. 예를 들어 Amazon EMR 릴리스 버전 5.18.0의 Hadoop 구성 요소에서 버전이 `2.8.4-amzn-1`입니다.

아티팩트 리포지토리 URL의 구문은 다음과 같습니다.

```
https://s3-endpoint/region-ID-emr-artifacts/emr-release-label/repos/maven/
```

- *`s3-endpoint`*는 리포지토리에 대한 리전의 Amazon Simple Storage Service(S3) 엔드포인트이고, *`region-ID`*는 해당 리전입니다. 예: `s3.us-west-1.amazonaws.com` 및 `us-west-1`. 자세한 내용은 Amazon Web Services 일반 참조에서 Amazon S3 엔드포인트를 참조하세요. 리전 사이의 아티팩트에는 차이가 없으므로 개발 환경에 가장 편리한 리전을 지정할 수 있습니다.
- *`emr-release-label`*은 코드를 실행할 Amazon EMR 클러스터의 릴리스 레이블입니다. 릴리스 레이블은 `emr-x.x.x` 형식을 따릅니다(예: `emr-5.36.2`). EMR 릴리스 시리즈 1개에는 다수의 릴리스가 포함될 수 있습니다. 예를 들어 EMR 릴리스 버전 5.24.1을 사용하고 있다면 아티팩트 리포지토리 URL에 5.24 시리즈의 첫 번째 EMR 릴리스 레이블인 `emr-5.24.0`을 사용하십시오.

  ```
  https://s3-endpoint/region-ID-emr-artifacts/emr-5.24.0/repos/maven/
  ```

Example Maven pom.xml을 위한 구성

아래의 pom.xml 예제는 us-west-1에 있는 아티팩트 리포지토리를 사용하여 emr-5.18.0 Apache Hadoop 및 Apache Hive 아티팩트에 대해 빌드하도록 Maven 프로젝트를 구성합니다. 아티팩트 리포

지토리에서 스냅샷 버전을 사용할 수 없으므로 pom.xml에서 스냅샷이 비활성화됩니다. 아래 예제의 줄임표(...)는 다른 구성 파라미터가 생략되었음을 나타내므로 Maven 프로젝트에 복사하지 마십시오.

```
<project>
 ...
 <repositories>
  ...
  <repository>
   <id>emr-5.18.0-artifacts</id>
   <name>EMR 5.18.0 Releases Repository</name>
   <releases>
    <enabled>true</enabled>
   </releases>
   <snapshots>
    <enabled>false</enabled>
   </snapshots>
   <url>https://s3.us-west-1.amazonaws.com/us-west-1-emr-artifacts/emr-5.18.0/repos/
maven/</url>
  </repository>
  ...
 </repositories>
 ...
 <dependencies>
  ...
  <dependency>
   <groupId>org.apache.hive</groupId>
   <artifactId>hive-exec</artifactId>
   <version>2.3.3-amzn-2</version>
  </dependency>
  <dependency>
   <groupId>org.apache.hadoop</groupId>
   <artifactId>hadoop-common</artifactId>
   <version>2.8.4-amzn-1</version>
  </dependency>
  ...
 </dependencies>

</project>
```

EMR 파일 시스템(EMRFS)

EMR 파일 시스템(EMRFS)은 모든 Amazon EMR 클러스터가 Amazon EMR에서 Amazon S3로 직접 일반 파일을 읽고 쓸 수 있도록 하는 HDFS 구현입니다. EMRFS는 Hadoop과 함께 사용하기 위해 Amazon S3에 영구 데이터를 저장하는 편리한 기능을 제공하면서 동시에 데이터 암호화 같은 기능도 제공합니다.

데이터 암호화는 EMRFS가 Amazon S3에 기록하는 객체를 암호화하여 EMRFS가 Amazon S3에서 암호화된 객체를 사용할 수 있도록 해줍니다. Amazon EMR 릴리스 버전 4.8.0 이상을 사용하고 있는 경우에는 보안 구성을 사용하여 다른 암호화 설정과 더불어 Amazon S3에서 EMRFS 객체에 대한 암호화를 설정할 수 있습니다. 자세한 내용은 암호화 옵션을 참조하세요. 이전 버전의 Amazon EMR을 사용하는 경우에는 수동으로 암호화 설정을 구성할 수 있습니다. 자세한 내용은 EMRFS 속성을 사용하여 Amazon S3 암호화 지정 단원을 참조하십시오.

Amazon S3는 모든 AWS 리전에서 모든 GET, PUT 및 LIST 작업에 대해 강력한 읽기 후 쓰기 일관성을 제공합니다. 즉, EMRFS를 사용하여 작성한 내용은 성능에 영향을 주지 ㅇ낳고 Amazon S3에서 읽을 수 있습니다. 자세한 내용은 Amazon S3 데이터 일관성 모델을 참조하십시오.

Amazon EMR 릴리스 버전 5.10.0 이상을 사용하는 경우에는 Amazon S3에서 EMRFS 데이터의 위치, 그룹 또는 클러스터 사용자에 따라 Amazon S3에 대한 EMRFS 요청에 다른 IAM 역할을 사용할 수 있습니다. 자세한 내용은 Amazon S3에 대한 EMRFS 요청의 IAM 역할 구성을 참조하세요.

> **Warning**
>
> Apache Spark 작업을 실행하는 Amazon EMR 클러스터의 추론적 실행을 활성화하기 전에 다음 정보를 검토합니다.
> EMRFS에는 EMRFS를 사용할 때 Amazon S3에 파일을 쓰는 데 최적화된 대체 OutputCommitter 구현인 EMRFS S3 최적화된 커미터가 포함되어 있습니다. EMRFS S3 최적화 커미터를 사용하지 않고 Amazon S3에 데이터를 쓰는 애플리케이션에서 Apache Spark 추론적 실행 기능을 켜면 SPARK-10063 설명에 설명된 데이터 정확성 문제가 발생할 수 있습니다. 이는 Amazon EMR 릴리스 5.19 이전의 Amazon EMR 버전을 사용하거나 ORC 및 CSV와 같은 형식으로 Amazon S3에 파일을 쓰는 경우 발생할 수 있습니다. 이러한 형식은 EMRFS S3 최적화 커미터에서 지원되지 않습니다. EMRFS S3 최적화된 커미터를 사용하기 위한 전체 요구 사항 목록은 EMRFS S3 최적화 커미터 요구 사항을 참조하세요.
> EMRFS 직접 쓰기는 일반적으로 다음과 같이 EMRFS S3 최적화 커미터가 지원되지 않는 경우에 사용됩니다.
>
> - Parquet 외 다른 출력 형식(예: ORC 또는 텍스트 등).

- Spark RDD API를 사용하는 Hadoop 파일.
- Hive Serde를 사용하는 Parquet. Hive metastore Parquet table conversion을 참조하세요.

다음 시나리오에서는 EMRFS 직접 쓰기가 사용되지 않습니다.

- EMRFS S3 최적화 커미터를 활성화하는 경우. EMRFS S3 최적화 커미터의 요구 사항을 참조하세요.
- partitionOverwriteMode를 dynamic으로 설정한 상태로 동적 파티션을 작성하는 경우
- 사용자 지정 파티션 위치에 쓸 때(예: Hive 기본 파티션 위치 규칙을 따르지 않는 위치).
- EMRFS 이외의 파일 시스템을 사용하는 경우(예: HDFS에 쓰기 또는 S3A 파일 시스템 사용).

애플리케이션이 Amazon EMR 5.14.0 이상에서 직접 쓰기를 사용하는지 여부를 확인하려면 Spark INFO 로깅을 활성화합니다. Spark 드라이버 로그나 Spark 실행기 컨테이너 로그에 'Direct Write: ENABLED'라는 텍스트가 포함된 로그 줄이 있는 경우 Spark 애플리케이션이 직접 쓰기를 사용하여 기록한 것입니다.
기본적으로 Amazon EMR 클러스터에서 추론적 실행은 OFF 상태입니다. 다음 두 가지 조건에 모두 해당하는 경우 추론적 실행을 켜지 않는 것이 좋습니다.

- Amazon S3에 데이터를 씁니다.
- 데이터는 Apache Parquet 이외의 형식 또는 EMRFS S3 최적화 커미터를 사용하지 않는 Apache Parquet 형식으로 작성됩니다.

Spark 추론적 실행을 켜고 EMRFS 직접 쓰기를 사용하여 Amazon S3에 데이터를 쓰는 경우 간헐적으로 데이터 손실이 발생할 수 있습니다. HDFS에 데이터를 쓰거나 EMRFS S3 최적화 커미터를 사용하여 Parquet에 데이터를 쓸 때 Amazon EMR은 직접 쓰기를 사용하지 않으므로 이 문제가 발생하지 않습니다.
Spark에서 Amazon S3로 직접 쓰는 EMRFS를 사용하는 형식으로 데이터를 쓰고 추론적 실행을 사용해야 하는 경우 HDFS에 쓴 다음, S3DistCP를 사용하여 출력 파일을 Amazon S3로 전송하는 것이 좋습니다.

주제

- 일관된 보기
- Amazon S3에서 EMRFS 데이터에 대한 액세스 권한 부여

- 기본 AWS Security Token Service 엔드포인트 관리
- EMRFS 속성을 사용하여 Amazon S3 암호화 지정

일관된 보기

> **Warning**
>
> 2023년 6월 1일에 EMRFS 일관된 보기가 향후 Amazon EMR 릴리스에 대한 표준 지원에서 종료됩니다. EMRFS 일관된 보기는 기존 릴리스에서도 계속 작동합니다.

2020년 12월 1일에 Amazon S3의 강력한 읽기 후 쓰기 일관성 기능이 출시됨에 따라 Amazon EMR 클러스터에서 EMRFS 일관된 보기(EMRFS CV)를 더 이상 사용할 필요가 없습니다. EMRFS CV는 Amazon EMR 클러스터가 Amazon S3 객체의 나열 및 쓰기 후 읽기 일관성을 확인할 수 있는 선택적 기능입니다. 클러스터를 생성하고 EMRFS CV를 켜면 Amazon EMR은 Amazon DynamoDB 데이터베이스를 생성하여 S3 객체의 목록 및 쓰기 후 읽기 일관성 추적에 사용하는 객체 메타데이터를 저장합니다. 이제 추가 비용이 발생하지 않도록 EMRFS CV를 끄고 CV에서 사용하는 DynamoDB 데이터베이스를 삭제할 수 있습니다. 다음 절차에서는 CV 기능을 확인하고, 기능을 끈 후에, 기능에서 사용하는 DynamoDB 데이터베이스를 삭제하는 방법을 설명합니다.

EMRFS CV 기능을 사용하고 있는지 확인하는 방법

1. 구성 탭으로 이동합니다. 클러스터의 구성이 다음과 같으면 EMRFS CV를 사용합니다.

   ```
   Classification=emrfs-site,Property=fs.s3.consistent,Value=true
   ```

2. 또는 AWS CLI 를 사용하여 `describe-cluster` API로 클러스터를 설명합니다. 출력에 `fs.s3.consistent: true`가 포함된 경우 클러스터는 EMRFS CV를 사용하는 것입니다.

Amazon EMR 클러스터에서 EMRFS CV를 끄는 방법

EMRFS CV 기능을 끄려면 다음 세 가지 옵션 중 하나를 사용합니다. 이러한 옵션을 프로덕션 환경에 적용하기 전에 테스트 환경에서 테스트해야 합니다.

1. 기존 클러스터를 중지하고 EMRFS CV 옵션을 사용하지 않고 새 클러스터를 시작하는 방법

   a. 클러스터를 중지하기 전에 데이터를 백업하고 사용자에게 알려야 합니다.

   b. 클러스터를 중지하려면 클러스터 종료 지침을 따릅니다.

c. Amazon EMR 콘솔을 사용하여 새 클러스터를 생성하는 경우 고급 옵션으로 이동합니다. 소프트웨어 설정 편집 섹션에서 EMRFS CV를 켜는 옵션을 선택 취소합니다. EMRFS 일관된 보기 확인란을 사용할 수 있는 경우 선택 취소한 상태로 유지합니다.

d. AWS CLI 를 사용하여 `create-cluster` API로 새 클러스터를 생성하는 경우 EMRFS CV를 켜는 `--emrfs` 옵션을 사용하지 마십시오.

e. SDK 또는를 사용하여 새 클러스터를 AWS CloudFormation 생성하는 경우 일관된 보기 구성에 나열된 구성을 사용하지 마십시오.

2. 클러스터를 복제하고 EMRFS CV를 제거하는 방법

a. Amazon EMR 콘솔에서 EMRFS CV를 사용하는 클러스터를 선택합니다.

b. 클러스터 세부 정보 페이지 상단에서 복제를 선택합니다.

c. 이전을 선택하고 1단계: 소프트웨어 및 단계로 이동합니다.

d. 소프트웨어 설정 편집에서 EMRFS CV를 제거합니다. 구성 편집에서 `emrfs-site` 분류에 있는 다음 구성을 삭제합니다. S3 버킷에서 JSON을 로드하는 경우 S3 객체를 수정해야 합니다.

```
[
 {"classification":
  "emrfs-site",
  "properties": {
   "fs.s3.consistent.retryPeriodSeconds":"10",
   "fs.s3.consistent":"true",
   "fs.s3.consistent.retryCount":"5",
   "fs.s3.consistent.metadata.tableName":"EmrFSMetadata"
  }
 }
]
```

3. 인스턴스 그룹을 사용하는 클러스터에서 EMRFS CV를 제거하는 방법

a. 다음 명령을 사용하여 단일 EMR 클러스터가 EMRFS CV와 연결된 DynamoDB 테이블을 사용하는지 또는 여러 클러스터가 테이블을 공유하는지 확인합니다. 테이블 이름은 일관된 보기 구성에 설명된 대로 `fs.s3.consistent.metadata.tableName`에서 지정됩니다. EMRFS CV에서 사용하는 기본 테이블 이름은 `EmrFSMetadata`입니다.

```
aws emr describe-cluster --cluster-id j-XXXXX | grep
 fs.s3.consistent.metadata.tableName
```

b. 클러스터가 DynamoDB 데이터베이스를 다른 클러스터와 공유하지 않는 경우 다음 명령을 사용하여 클러스터를 재구성하고 EMRFS CV를 비활성화합니다. 자세한 내용은 실행 중 클러스터에서 인스턴스 그룹 재구성을 참조하세요.

```
aws emr modify-instance-groups --cli-input-json file://disable-emrfs-1.json
```

이 명령은 수정할 파일을 엽니다. 다음 구성을 사용하여 파일을 수정합니다.

```
{
 "ClusterId": "j-xxxx",
 "InstanceGroups": [
  {
   "InstanceGroupId": "ig-xxxx",
   "Configurations": [
    {
     "Classification": "emrfs-site",
     "Properties": {
      "fs.s3.consistent": "false"
     },
     "Configurations": []
    }
   ]
  }
 ]
}
```

c. 클러스터가 DynamoDB 테이블을 다른 클러스터와 공유하는 경우, 공유 S3 위치의 객체를 수정하는 클러스터가 없는 시점에 모든 클러스터에서 EMRFS CV를 끕니다.

EMRFS CV와 연결된 Amazon DynamoDB 리소스를 삭제하는 방법

Amazon EMR 클러스터에서 EMRFS CV를 제거한 후에 EMRFS CV와 연결된 DynamoDB 리소스를 삭제합니다. 그렇게 하기 전까지는 EMRFS CV와 관련된 DynamoDB 요금이 계속 발생합니다.

1. DynamoDB 테이블의 CloudWatch 지표를 확인하고 해당 테이블이 클러스터에서 사용되지 않는지 확인합니다.
2. DynamoDB 테이블을 삭제합니다.

```
aws dynamodb delete-table --table-name <your-table-name>
```

EMRFS CV와 연결된 Amazon SQS 리소스를 삭제하는 방법

1. 불일치 알림을 Amazon SQS로 푸시하도록 클러스터를 구성한 경우 모든 SQS 대기열을 삭제할 수 있습니다.
2. 일관된 보기 구성에서 설명한 대로 `fs.s3.consistent.notification.SQS.queueName`에 지정된 Amazon SQS 대기열 이름을 찾습니다. 기본 대기열 이름 형식은 `EMRFS-Inconsistency-`*`<j-cluster ID>`*입니다.

```
aws sqs list-queues | grep 'EMRFS-Inconsistency'
aws sqs delete-queue -queue-url <your-queue-url>
```

EMRFS CLI 사용을 중지하는 방법

- EMRFS CLI는 EMRFS CV가 생성하는 메타데이터를 관리합니다. 향후 Amazon EMR 릴리스에서 EMRFS CV에 대한 표준 지원이 종료됨에 따라 EMRFS CLI에 대한 지원도 종료됩니다.

주제

일관된 보기 활성화

AWS Management Console, 또는 구성 분류를 사용하여 EMRFS에 대해 Amazon S3 서버 측 암호화 또는 일관된 보기를 활성화할 수 있습니다. AWS CLI `emrfs-site`

콘솔을 사용하여 일관성 보기를 구성하려면

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.

3. Step 1: Software and Steps(1단계: 소프트웨어 및 설정) 및 Step 2: Hardware(2단계: 하드웨어)에 대한 설정을 선택합니다.
4. Step 3: General Cluster Settings(3단계: 일반 클러스터 설정의 경우 추가 옵션에서 EMRFS consistent view(EMRFS 일관성 보기)를 선택합니다.
5. EMRFS Metadata store(EMRFS 메타데이터 스토어)의 경우 메타데이터 스토어 이름을 입력합니다. 기본값은 **EmrFSMetadata**입니다. EmrFSMetadata 테이블이 없는 경우 DynamoDB에서 자동으로 생성됩니다.

   > **Note**
   >
   > Amazon EMR은 클러스터가 종료될 때 DynamoDB에서 EMRFS 메타데이터를 자동으로 제거하지 않습니다.

6. 재시도 횟수에 정수 값을 입력합니다. 불일치가 감지되면 EMRFS에서는 이 횟수만큼 Amazon S3를 직접 호출하려고 시도합니다. 기본값은 **5**입니다.
7. Retry period (in seconds)(재시도 기간(초))의 경우 정수 값을 입력합니다. 이 값은 EMRFS에서 재시도 간에 기다리는 시간입니다. 기본값은 **10**입니다.

   > **Note**
   >
   > 후속 시도에는 지수 백오프가 사용됩니다.

를 사용하여 일관된 보기가 활성화된 클러스터를 시작하려면 AWS CLI

의 현재 버전을 설치하는 것이 좋습니다 AWS CLI. 최신 릴리스를 다운로드하려면 https://aws.amazon.com/cli/를 참조하세요.

- > **Note**
  >
  > 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --instance-type m5.xlarge --instance-count 3 --emrfs
 Consistent=true \
--release-label emr-7.9.0 --ec2-attributes KeyName=myKey
```

를 사용하여 일관된 보기가 활성화되어 있는지 확인하려면 AWS Management Console

- 일관성 보기가 콘솔에 활성화되었는지 여부를 확인하려면 Cluster List(클러스터 목록)로 이동하고 클러스터 세부 정보를 표시할 클러스터 이름을 선택합니다. "EMRFS consistent view" 필드에는 `Enabled` 또는 `Disabled` 값이 있습니다.

`emrfs-site.xml` 파일을 검사하여 일관성 보기가 활성화되었는지 여부를 확인하려면

- 클러스터의 마스터 노드에서 `emrfs-site.xml` 구성 파일을 검사하여 일관성이 활성화되었는지 여부를 확인할 수 있습니다. `fs.s3.consistent`의 부울 값이 `true`로 설정된 경우 Amazon S3와 관련된 파일 시스템 작업에 대해 일관된 보기가 활성화됩니다.

EMRFS의 일관된 보기가 Amazon S3에서 객체를 추적하는 방법 이해

EMRFS는 객체의 일관된 보기에 대한 정보를 EMRFS 메타데이터에 추가함으로써 Amazon S3에서 이러한 일관된 보기를 생성합니다. 다음의 경우 EMRFS는 이러한 목록을 메타데이터에 추가합니다.

- 객체가 Amazon EMR 작업 과정 중에 EMRFS에 의해 작성되는 경우
- EMRFS CLI를 사용하여 객체를 EMRFS 메타데이터와 동기화하거나 EMRFS 메타데이터로 가져오는 경우

EMRFS에서 읽은 객체는 메타데이터에 자동으로 추가되지 않습니다. EMRFS가 객체를 삭제하면 EMRFS CLI를 사용하여 목록을 삭제할 때까지 목록이 메타데이터에 삭제된 상태로 유지됩니다. CLI에 대한 자세한 내용은 EMRFS CLI 명령 참조 단원을 참조하십시오. EMRFS 메타데이터에서 목록 삭제에 대한 자세한 내용은 EMRFS 일관된 보기 메타데이터 단원을 참조하십시오.

Amazon S3 작업마다 EMRFS는 메타데이터에서 일관된 보기의 객체 세트에 대한 정보를 확인합니다. 이러한 작업 중 하나를 수행하는 동안 Amazon S3가 일관되지 않음이 확인되는 경우 EMRFS는 `emrfs-site` 구성 속성에 정의된 파라미터에 따라 작업을 재시도합니다. 재시도 횟수가 모두 사용된 후에는 `ConsistencyException`을 발생하거나 예외를 기록하고 워크플로를 계속 진행합니다. 재시도 로직에 대한 자세한 내용은 재시도 로직 단원을 참조하십시오. 예를 들면 로그에 `ConsistencyExceptions`가 있을 수 있습니다.

- listStatus: 메타데이터 항목 `/S3_bucket/dir/object`에 해당하는 Amazon S3 객체 없음
- getFileStatus: `dir/file` 키가 메타데이터에 있지만 Amazon S3에 없음

EMRFS의 일관된 보기에서 추적되는 객체를 Amazon S3에서 직접 삭제하더라도 Amazon S3의 메타데이터에 존재하는 항목으로 여전히 나열되므로 EMRFS는 해당 객체를 일관되지 않은 항목을 간주합니다. 메타데이터가 EMRFS가 Amazon S3에서 추적하는 객체와 동기화되지 않은 경우 EMRFS CLI에서 sync 하위 명령을 사용하여 Amazon S3를 반영하도록 메타데이터를 재설정할 수 있습니다. 메타데이터와 Amazon S3 간의 불일치를 찾기 위해 diff를 사용합니다. 마지막으로 EMRFS는 메타데이터에서 참조되는 객체의 일관된 보기만 제공합니다. 동일한 Amazon S3 경로에 추적되지 않는 다른 객체가 있을 수도 있습니다. EMRFS는 Amazon S3 경로에서 객체를 나열할 때 메타데이터에서 추적되는 객체와 Amazon S3 경로에서 추적되는 객체의 상위 세트를 반환합니다.

재시도 로직

EMRFS는 특정 재시도 횟수 동안 메타데이터에서 추적되는 객체에 대한 목록 일관성 확인을 시도합니다. 기본값은 5입니다. 재시도 횟수를 초과하는 경우 `fs.s3.consistent.throwExceptionOnInconsistency`가 `false`로 설정되지 않았으면 원래 작업에서 실패 메시지를 반환합니다. 이 옵션이 false이면 객체가 일관되지 않은 항목으로만 기록됩니다. EMRFS는 기본적으로 지수 백오프 재시도 정책을 사용하지만 고정 정책으로 설정할 수도 있습니다. 또한 사용자는 예외 발생 없이 작업의 나머지 부분을 지속하기 전에 특정 시간 동안 재시도하려고 할 수도 있습니다. `fs.s3.consistent.throwExceptionOnInconsistency`를 `false`로, `fs.s3.consistent.retryPolicyType` 을 `fixed`로, 그리고 `fs.s3.consistent.retryPeriodSeconds`를 원하는 값으로 설정하면 가능합니다. 다음 예제에서는 일관성이 활성화된 상태로 클러스터를 생성합니다. 이 경우, 불일치 항목을 기록하고 10초의 고정 재시도 간격을 설정합니다.

Example 재시도 기간을 고정 값으로 설정

```
aws emr create-cluster --release-label emr-7.9.0 \
--instance-type m5.xlarge --instance-count 1 \
--emrfs Consistent=true,Args=[fs.s3.consistent.throwExceptionOnInconsistency=false,
 fs.s3.consistent.retryPolicyType=fixed,fs.s3.consistent.retryPeriodSeconds=10] --ec2-
attributes KeyName=myKey
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

자세한 내용은 [일관된 보기](#) 단원을 참조하십시오.

IMDS의 EMRFS 구성은 리전 직접 호출을 가져옴

EMRFS는 IMDS(인스턴스 메타데이터 서비스)를 사용하여 인스턴스 리전과 Amazon S3, DynamoDB 또는 AWS KMS 엔드포인트를 가져옵니다. 하지만 IMDS는 처리할 수 있는 요청 수에 제한이 있으며, 이 제한을 초과하는 요청은 실패합니다. 이 IMDS 제한으로 인해 EMRFS 초기화에 실패하고 쿼리 또는 명령이 실패할 수 있습니다. 다음과 같은 무작위 지수 백오프 재시도 메커니즘과 emrfs-site.xml의 폴백 리전 구성 속성을 사용하여 모든 재시도가 실패하는 시나리오를 해결할 수 있습니다.

```
<property>
    <name>fs.s3.region.retryCount</name>
    <value>3</value>
    <description>
    Maximum retries that would be attempted to get AWS region.
    </description>
</property>
<property>
    <name>fs.s3.region.retryPeriodSeconds</name>
    <value>3</value>
    <description>
    Base sleep time in second for each get-region retry.
    </description>
</property>
<property>
    <name>fs.s3.region.fallback</name>
    <value>us-east-1</value>
    <description>
    Fallback to this region after maximum retries for getting AWS region have been
 reached.
    </description>
</property>
```

EMRFS 일관된 보기 메타데이터

EMRFS 일관된 보기는 DynamoDB 테이블을 사용하여 Amazon S3에서 EMRFS와 동기화되었거나 EMRFS에서 생성된 객체를 추적하는 방법으로 일관성을 추적합니다. 이 메타데이터는 모든 작업(읽기, 쓰기, 업데이트 및 복사)를 추적하는 데 사용되며, 실제 내용은 저장되지 않습니다. 이 메타데이터는 Amazon S3에서 수신된 객체나 메타데이터가 예상과 일치하는지 여부를 검증하는 데 사용됩니다. 이 확인을 통해 EMRFS는 EMRFS에서 Amazon S3에 기록하는 새 객체와 EMRFS와 동기화된 객체에 대한 쓰기 후 읽기 일관성 및 목록 일관성을 검사할 수 있습니다. 여러 클러스터가 동일한 메타데이터를 공유할 수 있습니다.

메타데이터에 항목을 추가하는 방법

`sync` 또는 `import` 하위 명령을 사용하여 항목을 메타데이터에 추가할 수 있습니다. `sync`는 경로에 Amazon S3 객체의 상태를 반영하지만 `import`는 새 항목을 메타데이터에 추가하는 데에만 사용됩니다. 자세한 내용은 EMRFS CLI 명령 참조 단원을 참조하십시오.

메타데이터와 Amazon S3 객체 간 차이를 확인하는 방법

메타데이터와 Amazon S3 간의 차이를 확인하려면 EMRFS CLI의 `diff` 하위 명령을 사용합니다. 자세한 내용은 EMRFS CLI 명령 참조 단원을 참조하십시오.

메타데이터 작업이 제한되는지 여부를 알아보는 방법

EMRFS는 메타데이터에 대한 읽기 및 쓰기 작업의 기본 처리량 제한을 각각 500 유닛 및 100 유닛으로 설정합니다. 객체 또는 버킷 수가 많으면 작업이 이 용량을 초과하여 DynamoDB에서 작업을 제한합니다. 예를 들어 애플리케이션에서 이러한 용량 제한을 초과하는 작업을 수행하려는 경우 EMRFS에서 `ProvisionedThroughputExceededException`이 발생할 수 있습니다. 제한 시 EMRFS CLI 도구는 작업이 완료될 때까지 또는 Amazon EMR에서 Amazon S3로 객체 쓰기 최대 재시도 값에 도달할 때 지수 백오프를 사용하여 DynamoDB 테이블에 쓰는 작업을 재시도합니다.

처리 용량 한도를 구성할 수 있습니다. 그러나 DynamoDB의 읽기 및 쓰기 작업에 대한 파티션 제한은 초당 3,000의 읽기 용량 유닛(RCU) 및 1,000의 쓰기 용량 유닛(WCU)이 적용됩니다. 제한으로 인한 `sync` 실패를 방지하려면 읽기 작업의 처리량을 3,000 RCU, 쓰기 작업의 처리량을 1,000 WCU 미만으로 제한하는 것이 좋습니다. 사용자 지정 처리량 용량 제한을 설정하는 방법에 대한 지침은 일관된 보기 구성 섹션을 참조하세요.

제한된 읽기 및 쓰기 요청 수를 볼 수 있는 DynamoDB 콘솔에서 EMRFS 메타데이터에 대한 Amazon CloudWatch 지표도 볼 수 있습니다. 제한된 요청에 대해 0이 아닌 값을 지정한 읽기 또는 쓰기 작업에 대해 할당된 처리 용량을 높이는 것이 애플리케이션에 도움이 될 수 있습니다. 또한 애플리케이션이 장시간 읽기 또는 쓰기 작업에 대해 할당된 최대 처리 용량에 근접하는 경우에도 성능 이점을 얻을 수 있습니다.

주목할 만한 EMRFS 작업에 대한 처리량 특성

읽기 및 쓰기 작업의 처리 용량 기본값은 각각 400단위 및 100단위입니다. 다음과 같은 성능 특성은 특정 작업에 필요한 처리량을 결정하는 데 도움이 됩니다. 이러한 테스트는 단일 노드 `m3.large` 클러스터를 사용하여 수행되었습니다. 모든 작업이 단일 스레드입니다. 성능은 특정 애플리케이션 특성에 따라 크게 달라지므로 파일 시스템 작업을 최적화하기 위해 실험이 필요할 수도 있습니다.

Operation	초당 평균 읽기	초당 평균 쓰기
create(객체)	26.79	6.70
delete(객체)	10.79	10.79
delete(1000개 객체를 포함하는 디렉터리)	21.79	338.40
getFileStatus(객체)	34.70	0
getFileStatus(디렉터리)	19.96	0
listStatus(객체 1개를 포함하는 디렉터리)	43.31	0
listStatus(10개 객체를 포함하는 디렉터리)	44.34	0
listStatus(100개 객체를 포함하는 디렉터리)	84.44	0
listStatus(1,000개 객체를 포함하는 디렉터리)	308.81	0
listStatus(10,000개 객체를 포함하는 디렉터리)	416.05	0
listStatus(100,000개 객체를 포함하는 디렉터리)	823.56	0
listStatus(1M 객체를 포함하는 디렉터리)	882.36	0
mkdir(120초 동안 지속)	24.18	4.03
mkdir	12.59	0
rename(객체)	19.53	4.88

Operation	초당 평균 읽기	초당 평균 쓰기
rename(1000개 객체를 포함하는 디렉터리)	23.22	339.34

메타데이터 스토어에서 오래된 데이터를 삭제하는 단계를 제출하려면

사용자가 DynamoDB 기반 메타데이터에서 특정 항목을 제거해야 할 때가 있습니다. 이 작업을 수행하면 테이블 관련 스토리지 비용을 절감할 수 있습니다. 사용자는 EMRFS CLI `delete` 하위 명령을 사용하여 특정 항목을 수동으로 또는 프로그래밍 방식으로 삭제할 수 있습니다. 하지만 메타데이터의 항목을 삭제할 경우 EMRFS에서 더 이상 일관성 검사가 수행되지 않습니다.

작업 완료 후 프로그래밍 방식으로 삭제하는 작업은 EMRFS CLI에서 명령을 실행하는 최종 단계를 클러스터에 제출함으로써 수행할 수 있습니다. 예를 들어 다음 명령을 입력하여 클러스터에 2일 이상 지난 모든 항목을 삭제하도록 하는 단계를 클러스터에 제출합니다.

```
aws emr add-steps --cluster-id j-2AL4XXXXXX5T9 --steps Name="emrfsCLI",Jar="command-
runner.jar",Args=["emrfs","delete","--time","2","--time-unit","days"]
{
    "StepIds": [
        "s-B12345678902"
    ]
}
```

반환된 StepId 값을 사용하여 작업 결과에 대한 로그를 확인합니다.

CloudWatch 및 Amazon SQS에 대한 일관성 알림 구성

EMRFS에서 Amazon S3 최종 일관성 문제와 관련하여 CloudWatch 지표 및 Amazon SQS 메시지를 활성화할 수 있습니다.

CloudWatch

CloudWatch 지표가 활성화되면 Amazon S3의 최종 일관성으로 인해 `FileSystem` API 직접 호출이 실패할 때마다 불일치라는 이름의 지표가 푸시됩니다.

Amazon S3 최종 일관성 문제에 대해 CloudWatch 지표를 보는 방법

CloudWatch 콘솔에서 불일치 지표를 보려면 EMRFS 지표를 선택한 다음 JobFlowID 및 지표 이름 페어를 선택합니다. 예: `j-162XXXXXXM2CU ListStatus`, `j-162XXXXXXM2CU GetFileStatus` 등.

1. https://console.aws.amazon.com/cloudwatch/에서 CloudWatch 콘솔을 엽니다.
2. [Dashboard]의 [Metrics] 섹션에서 [EMRFS]를 선택합니다.
3. [Job Flow Metrics] 창에서 하나 이상의 [JobFlowId/Metric Name] 페어를 선택합니다. 측정치를 나타내는 그림이 아래 창에 나타납니다.

Amazon SQS

Amazon SQS 알림이 활성화되면 EMRFS가 초기화될 때 이름이 `EMRFS-Inconsistency-<jobFlowId>`인 Amazon SQS 대기열이 생성됩니다. Amazon S3의 최종 일관성으로 인해 `FileSystem` API 직접 호출에 실패하면 Amazon SQS 메시지가 대기열로 푸시됩니다. 이 메시지에는 JobFlowId, API, 일관되지 않은 경로 목록, 스택 추적 등의 정보가 포함됩니다. Amazon SQS 콘솔이나 EMRFS `read-sqs` 명령을 사용하여 메시지를 읽을 수 있습니다.

Amazon S3 최종 일관성 문제에 대한 Amazon SQS 메시지를 관리하는 방법

Amazon S3 최종 일관성 문제에 대한 Amazon SQS 메시지는 EMRFS CLI를 사용하여 읽을 수 있습니다. EMRFS Amazon SQS 대기열에서 메시지를 읽으려면 `read-sqs` 명령을 입력하고 결과 출력 파일에 대해 프라이머리 노드의 로컬 파일 시스템에 출력 위치를 지정합니다.

또한 `delete-sqs` 명령을 사용하여 EMRFS Amazon SQS 대기열을 삭제할 수도 있습니다.

1. Amazon SQS 대기열의 메시지를 읽으려면 다음 명령을 입력합니다. *queuename*을 구성된 Amazon SQS 대기열의 이름으로 바꾸고 */path/filename*을 출력 파일의 경로로 바꿉니다.

```
emrfs read-sqs --queue-name queuename --output-file /path/filename
```

   예를 들어 기본 대기열에서 Amazon SQS 메시지를 읽고 출력하려면 다음을 입력합니다.

```
emrfs read-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU --output-file /
path/filename
```

> **Note**
>
> 또한 각각 `-q` 및 `-o` 대신에 `--queue-name` 및 `--output-file` 바로 가기를 사용할 수도 있습니다.

2. Amazon SQS 대기열을 삭제하려면 다음 명령을 입력합니다.

```
emrfs delete-sqs --queue-name queuename
```

예를 들어 기본 대기열을 삭제하려면 다음을 입력합니다.

```
emrfs delete-sqs --queue-name EMRFS-Inconsistency-j-162XXXXXXM2CU
```

> **Note**
>
> 또한 `-q` 대신에 `--queue-name` 바로 가기를 사용할 수도 있습니다.

일관된 보기 구성

`emrfs-site` 속성에서 구성 속성을 통해 추가 설정을 제공함으로써 일관성 보기에 대해 추가 설정을 구성할 수 있습니다. 예를 들어 emrfs-site 구성 분류(Amazon EMR 릴리스 버전 4.x 이상만)나 부트스트랩 작업을 사용하여 프라이머리 노드에서 emrfs-site.xml 파일을 구성함으로써 CLI `--emrfs` 옵션에 다음과 같은 인수를 제공해 다른 기본 DynamoDB 처리량을 선택할 수 있습니다.

Example 클러스터 시작 시 기본 메타데이터 읽기 및 쓰기 값 변경

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--emrfs Consistent=true,Args=[fs.s3.consistent.metadata.read.capacity=600,\
fs.s3.consistent.metadata.write.capacity=300] --ec2-attributes KeyName=myKey
```

또는 다음과 같은 구성 파일을 사용하고 이 파일을 로컬로 저장하거나 Amazon S3에 저장합니다.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.consistent.metadata.read.capacity": "600",
        "fs.s3.consistent.metadata.write.capacity": "300"
      }
    }
 ]
```

다음 구문으로 생성된 구성을 사용합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hive \
```

```
--instance-type m5.xlarge --instance-count 2 --configurations file://./myConfig.json
```

> **Note**
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

구성 또는 AWS CLI `--emrfs` 인수를 사용하여 다음 옵션을 설정할 수 있습니다. 이러한 인수에 대한 자세한 내용은 [AWS CLI 명령 참조](#)를 참조하세요.

일관된 보기에 대한 **emrfs-site.xml** 속성

속성	기본값	설명
`fs.s3.consistent`	**false**	**true**로 설정할 경우 이 속성은 DynamoDB를 사용하여 일관성을 제공하도록 EMRFS를 구성합니다.
`fs.s3.consistent.retryPolicyType`	**exponential**	이 속성은 일관성 문제로 재시도 시 사용할 정책을 식별합니다. exponential, fixed 또는 none을 옵션으로 사용할 수 있습니다.
`fs.s3.consistent.retryPeriodSeconds`	**1**	이 속성은 일관성 재시도 사이 대기 시간을 설정합니다.
`fs.s3.consistent.retryCount`	**10**	이 속성은 비일관성이 검색될 때 최대 재시도 횟수를 설정합니다.
`fs.s3.consistent.throwExceptionOnInconsistency`	**true**	이 속성은 일관성 예외를 발생할지 아니면 기록할지 여부를 결정합니다. **true**로 설정할 경우 `ConsistencyException` 이 발생됩니다.

속성	기본값	설명
`fs.s3.consistent.metadata.autoCreate`	**`true`**	**`true`**로 설정할 경우 이 속성은 메타데이터 테이블의 자동 생성을 활성화합니다.
`fs.s3.consistent.metadata.etag.verification.enabled`	**`true`**	Amazon EMR 5.29.0의 경우 이 속성은 기본적으로 활성화되어 있습니다. 활성화하면 EMRFS는 S3 ETag를 사용하여 읽고 있는 객체가 사용 가능한 최신 버전인지 확인합니다. 이 기능은 동일한 이름을 유지하면서 S3의 파일을 덮어쓰는 업데이트 후 읽기 사용 사례에 유용합니다. 이 ETag 확인 기능은 현재 S3 Select에서 작동하지 않습니다.
`fs.s3.consistent.metadata.tableName`	**`EmrFSMetadata`**	이 속성은 DynamoDB에서 메타데이터 테이블의 이름을 지정합니다.
`fs.s3.consistent.metadata.read.capacity`	**`500`**	이 속성은 메타데이터 테이블이 생성될 때 프로비저닝할 DynamoDB 읽기 용량을 지정합니다.
`fs.s3.consistent.metadata.write.capacity`	**`100`**	이 속성은 메타데이터 테이블이 생성될 때 프로비저닝할 DynamoDB 쓰기 용량을 지정합니다.
`fs.s3.consistent.fastList`	**`true`**	**`true`**로 설정할 경우 이 속성은 여러 스레드를 사용하여 디렉터리를 나열합니다(필요한 경우). 이 속성을 사용하려면 일관성이 활성화되어야 합니다.

속성	기본값	설명
fs.s3.consistent.fastList.prefetchMetadata	**false**	**true**로 설정할 경우 이 속성은 20,000개 이상의 항목을 포함하는 디렉터리에 대해 메타데이터 미리 가져오기 기능을 활성화합니다.
fs.s3.consistent.notification.CloudWatch	**false**	**true**로 설정할 경우 Amazon S3 최종 일관성 문제로 인해 실패하는 FileSystem API 직접 호출에 대해 CloudWatch 지표가 활성화됩니다.
fs.s3.consistent.notification.SQS	**false**	**true**로 설정할 경우 최종 일관성 알림이 Amazon SQS 대기열로 푸시됩니다.
fs.s3.consistent.notification.SQS.queueName	**EMRFS-Inconsistency-<jobFlowId>**	이 속성을 변경하면 Amazon S3 최종 일관성 문제에 관한 메시지에 대해 사용자 고유의 SQS 대기열 이름을 지정할 수 있습니다.
fs.s3.consistent.notification.SQS.customMsg	**none**	이 속성을 통해 Amazon S3 최종 일관성 문제에 관한 SQS 메시지에 포함된 사용자 지정 정보를 지정할 수 있습니다. 이 속성에 대한 값을 지정하지 않은 경우 메시지의 해당 필드가 비어 있습니다.
fs.s3.consistent.dynamodb.endpoint	**none**	이 속성을 사용하면 일관성 보기 메타데이터에 대한 사용자 지정 DynamoDB 엔드포인트를 지정할 수 있습니다.

속성	기본값	설명
`fs.s3.useRequesterPaysHeader`	**`false`**	**`true`**로 설정하면 이 속성을 사용하여 요청 지급인 옵션이 활성화된 상태에서 Amazon S3의 버킷 요청이 허용됩니다.

EMRFS CLI 명령 참조

EMRFS CLI는 Amazon EMR 릴리스 버전 3.2.1 이상을 사용하여 생성된 모든 클러스터 프라이머리 노드에 기본적으로 설치됩니다. EMRFS CLI를 사용하여 일관성 보기에 대한 메타데이터를 관리할 수 있습니다.

> **Note**
>
> emrfs 명령은 VT100 터미널 에뮬레이션에만 지원됩니다. 하지만 다른 터미널 에뮬레이터 모드에서 작동할 수도 있습니다.

emrfs top-level 명령

emrfs 최상위 명령은 다음 구조를 지원합니다.

```
emrfs [describe-metadata | set-metadata-capacity | delete-metadata | create-metadata | \
list-metadata-stores | diff | delete | sync | import ] [options] [arguments]
```

다음 표의 설명에 따라 [arguments] 유무에 상관없이 [options]를 지정합니다. 하위 명령(`describe-metadata`, `set-metadata-capacity` 등)에 대한 [options]는 아래 하위 명령을 참조하십시오.

emrfs의 [options]

옵션	설명	필수
	Amazon S3에 객체를 쓰고 DynamoDB에서 메타데이터 스토어를 생성하거나 액세스하는 데 사용하는 AWS 액세스 키입니다. 기본적으로 *AWS_ACCES*	아니요

옵션	설명	필수
-a *AWS_ACCESS_KEY_ID* | --access-key *AWS_ACCESS_KEY_ID*	*S_KEY_ID* 는 클러스터를 생성하는 데 사용되는 액세스 키로 설정됩니다.	
-s *AWS_SECRET_ACCESS_KEY* | --secret-key *AWS_SECRET_ACCESS_KEY*	Amazon S3에 객체를 쓰고 DynamoDB에서 메타데이터 스토어를 생성하거나 액세스하는 데 사용하는 액세스 키와 연결된 AWS 보안 키입니다. 기본적으로 *AWS_SECRET_ACCESS_KEY* 는 클러스터를 생성하는 데 사용되는 액세스 키와 연결된 비밀 키로 설정됩니다.	아니요
-v | --verbose	출력을 상세 표시 모드로 설정합니다.	아니요
-h | --help	emrfs 명령에 대한 도움말 메시지와 사용법 설명을 표시합니다.	아니요

emrfs describe-metadata 하위 명령

emrfs describe-metadata의 [options]

옵션	설명	필수
-m *METADATA_NAME* | --metadata-name *METADATA_NAME*	*METADATA_NAME* 은 DynamoDB 메타데이터 테이블 이름입니다. *METADATA_NAME* 인수를 지정하지 않을 경우 기본값은 `EmrFSMetadata` 입니다.	아니요

Example emrfs describe-metadata 예제

다음 예제에서는 기본 메타데이터 테이블을 설명합니다.

```
$ emrfs describe-metadata
EmrFSMetadata
  read-capacity: 400
```

```
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 12
```

emrfs set-metadata-capacity 하위 명령

emrfs set-metadata-capacity의 [options]

옵션	설명	필수
-m *METADATA_NAME* | --metadata-name *METADATA_NAME*	*METADATA_NAME* 은 DynamoDB 메타데이터 테이블 이름입니다. *METADATA_NAME* 인수를 지정하지 않을 경우 기본값은 EmrFSMetadata 입니다.	아니요
-r *READ_CAPACITY* | --read-capacity *READ_CAPACITY*	메타데이터 테이블에 대해 요청한 읽기 처리 용량입니다. *READ_CAPACITY* 인수를 지정하지 않을 경우 기본값은 400입니다.	아니요
-w *WRITE_CAPACITY* | --write-capacity *WRITE_CAPACITY*	메타데이터 테이블에 대해 요청한 쓰기 처리 용량입니다. *WRITE_CAPACITY* 인수를 지정하지 않을 경우 기본값은 100입니다.	아니요

Example emrfs set-metadata-capacity 예제

다음 예제에서는 메타데이터 테이블 600에 대해 읽기 처리 용량을 150으로 설정하고 쓰기 용량을 EmrMetadataAlt으로 설정합니다.

```
$ emrfs set-metadata-capacity --metadata-name EmrMetadataAlt  --read-capacity 600 --
write-capacity 150
  read-capacity: 400
  write-capacity: 100
  status: UPDATING
  approximate-item-count (6 hour delay): 0
```

emrfs delete-metadata 하위 명령

emrfs delete-metadata의 [options]

옵션	설명	필수
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* 은 DynamoDB 메타데이터 테이블 이름입니다. *METADATA_NAME* 인수를 지정하지 않을 경우 기본값은 `EmrFSMetadata` 입니다.	아니요

Example emrfs delete-metadata 예제

다음 예제에서는 기본 메타데이터 테이블을 삭제합니다.

```
$ emrfs delete-metadata
```

emrfs create-metadata 하위 명령

emrfs create-metadata의 [options]

옵션	설명	필수
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* 은 DynamoDB 메타데이터 테이블 이름입니다. *METADATA_NAME* 인수를 지정하지 않을 경우 기본값은 `EmrFSMetadata` 입니다.	아니요
`-r` *READ_CAPACITY* | `--read-capacity` *READ_CAPACITY*	메타데이터 테이블에 대해 요청한 읽기 처리 용량입니다. *READ_CAPACITY* 인수를 지정하지 않을 경우 기본값은 `400`입니다.	아니요
`-w` *WRITE_CAPACITY* | `--write-capacity` *WRITE_CAPACITY*	메타데이터 테이블에 대해 요청한 쓰기 처리 용량입니다. *WRITE_CAPACITY* 인수를 지정하지 않을 경우 기본값은 `100`입니다.	아니요

Example emrfs create-metadata 예제

다음 예제에서는 메타데이터 테이블 `EmrFSMetadataAlt`를 생성합니다.

```
$ emrfs create-metadata -m EmrFSMetadataAlt
Creating metadata: EmrFSMetadataAlt
EmrFSMetadataAlt
  read-capacity: 400
  write-capacity: 100
  status: ACTIVE
  approximate-item-count (6 hour delay): 0
```

emrfs list-metadata-stores 하위 명령

emrfs list-metadata-stores 하위 명령에는 [options]가 없습니다.

Example List-metadata-stores 예제

다음 예제는 메타데이터 테이블을 나열합니다.

```
$ emrfs list-metadata-stores
  EmrFSMetadata
```

emrfs diff 하위 명령

emrfs diff의 [options]

옵션	설명	필수
`-m` *`METADATA_NAME`* | `--metadata-name` *`METADATA_NAME`*	*`METADATA_NAME`* 은 DynamoDB 메타데이터 테이블 이름입니다. *`METADATA_NAME`* 인수를 지정하지 않을 경우 기본값은 `EmrFSMetadata` 입니다.	아니요
`s3://s3Path`	메타데이터 테이블과 비교하기 위한 Amazon S3 버킷 경로. 버킷을 재귀적으로 동기화합니다.	예

Example emrfs diff 예제

다음 예제에서는 기본 메타데이터 테이블을 Amazon S3 버킷과 비교합니다.

```
$ emrfs diff s3://elasticmapreduce/samples/cloudfront
BOTH | MANIFEST ONLY | S3 ONLY
DIR elasticmapreduce/samples/cloudfront
DIR elasticmapreduce/samples/cloudfront/code/
DIR elasticmapreduce/samples/cloudfront/input/
DIR elasticmapreduce/samples/cloudfront/logprocessor.jar
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-14.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-15.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-16.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-17.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-18.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-19.WxYz1234
DIR elasticmapreduce/samples/cloudfront/input/XABCD12345678.2009-05-05-20.WxYz1234
DIR elasticmapreduce/samples/cloudfront/code/cloudfront-loganalyzer.tgz
```

emrfs delete 하위 명령

emrfs delete의 [options]

옵션	설명	필수
`-m METADATA_NAME | --metadata-name METADATA_NAME`	*METADATA_NAME* 은 DynamoDB 메타데이터 테이블 이름입니다. *METADATA_NAME* 인수를 지정하지 않을 경우 기본값은 `EmrFSMetadata` 입니다.	아니요
s3://s3Path	일관된 보기에 대해 추적 중인 Amazon S3 버킷의 경로입니다. 버킷을 재귀적으로 동기화합니다.	예
`-t TIME | --time TIME`	만료 시간(시간 단위 인수를 사용하여 해석됨). 지정된 버킷에 대해 *TIME* 인수보다 오래된 모든 메타데이터 항목이 삭제됩니다.	
`-u UNIT | --time-unit UNIT`	시간 인수를 해석하는 데 사용되는 단위(nanoseconds, microseconds, milliseconds, seconds, minutes, hours 또는 days)입니다. 인수를 지정하지 않을 경우 기본값은 `days`입니다.	

옵션	설명	필수
`--read-consumption` *`READ_CONSUMPTION`*	delete 작업에 사용할 요청된 사용 가능한 읽기 처리량입니다. *`READ_CONSUMPTION`* 인수를 지정하지 않는 경우 기본값은 `400`입니다.	아니요
`--write-consumption` *`WRITE_CONSUMPTION`*	delete 작업에 사용할 요청된 사용 가능한 쓰기 처리량입니다. *`WRITE_CONSUMPTION`* 인수를 지정하지 않는 경우 기본값은 `100`입니다.	아니요

Example emrfs delete 예제

다음 예제에서는 일관된 보기에 대해 추적 중인 메타데이터에서 Amazon S3 버킷의 모든 객체를 제거합니다.

```
$ emrfs delete s3://elasticmapreduce/samples/cloudfront
entries deleted: 11
```

emrfs import 하위 명령

emrfs import의 [options]

옵션	설명	필수
`-m` *`METADATA_NAME`* | `--metadata-name` *`METADATA_NAME`*	*`METADATA_NAME`* 은 DynamoDB 메타데이터 테이블 이름입니다. *`METADATA_NAME`* 인수를 지정하지 않을 경우 기본값은 `EmrFSMetadata` 입니다.	아니요
`s3://s3Path`	일관된 보기에 대해 추적 중인 Amazon S3 버킷의 경로입니다. 버킷을 재귀적으로 동기화합니다.	예
`--read-consumption` *`READ_CONSUMPTION`*	delete 작업에 사용할 요청된 사용 가능한 읽기 처리량입니다. *`READ_CONSUMPTION`* 인수를 지정하지 않는 경우 기본값은 `400`입니다.	아니요

옵션	설명	필수
`--write-consumption` *WRITE_CONSUMPTION*	delete 작업에 사용할 요청된 사용 가능한 쓰기 처리량입니다. *WRITE_CONSUMPTION* 인수를 지정하지 않는 경우 기본값은 `100`입니다.	아니요

Example emrfs import 예제

다음 예제에서는 일관된 보기에 대해 추적 중인 메타데이터와 함께 Amazon S3 버킷의 모든 객체를 가져옵니다. 알 수 없는 모든 키가 무시됩니다.

```
$ emrfs import s3://elasticmapreduce/samples/cloudfront
```

emrfs sync 하위 명령

emrfs sync의 [options]

옵션	설명	필수
`-m` *METADATA_NAME* | `--metadata-name` *METADATA_NAME*	*METADATA_NAME* 은 DynamoDB 메타데이터 테이블 이름입니다. *METADATA_NAME* 인수를 지정하지 않을 경우 기본값은 `EmrFSMetadata` 입니다.	아니요
s3://s3Path	일관된 보기에 대해 추적 중인 Amazon S3 버킷의 경로입니다. 버킷을 재귀적으로 동기화합니다.	예
`--read-consumption` *READ_CONSUMPTION*	delete 작업에 사용할 요청된 사용 가능한 읽기 처리량입니다. *READ_CONSUMPTION* 인수를 지정하지 않는 경우 기본값은 `400`입니다.	아니요
`--write-consumption` *WRITE_CONSUMPTION*	delete 작업에 사용할 요청된 사용 가능한 쓰기 처리량입니다. *WRITE_CONSUMPTION* 인수를 지정하지 않는 경우 기본값은 `100`입니다.	아니요

Example emrfs sync 명령 예제

다음 예제에서는 일관된 보기에 대해 추적 중인 메타데이터와 함께 Amazon S3 버킷의 모든 객체를 가져옵니다. 알 수 없는 모든 키가 삭제됩니다.

```
$ emrfs sync s3://elasticmapreduce/samples/cloudfront
Synching samples/cloudfront                                       0 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/code/                                 1 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/                                      2 added | 0 updated |
 0 removed | 0 unchanged
Synching samples/cloudfront/input/                                9 added | 0 updated |
 0 removed | 0 unchanged
Done synching s3://elasticmapreduce/samples/cloudfront            9 added | 0 updated |
 1 removed | 0 unchanged
creating 3 folder key(s)
folders written: 3
```

emrfs read-sqs 하위 명령

emrfs read-sqs의 [options]

옵션	설명	필수
`-q` *`QUEUE_NAME`* | `--queue-name` *`QUEUE_NAME`*	*`QUEUE_NAME`* 은 `emrfs-site.xml` 에 구성된 Amazon SQS 대기열 이름입니다. 기본값은 **`EMRFS-Inconsistency-<jobFlowId>`** 입니다.	예
`-o` *`OUTPUT_FILE`* | `--output-file` *`OUTPUT_FILE`*	*`OUTPUT_FILE`* 은 마스터 노드 로컬 파일 시스템의 출력 파일 경로입니다. 대기열에서 읽은 메시지가 이 파일에 기록됩니다.	예

emrfs delete-sqs 하위 명령

emrfs delete-sqs의 [options]

옵션	설명	필수

옵션	설명	필수
-q *QUEUE_NAME* | --queue-name *QUEUE_NAME*	*QUEUE_NAME* 은 emrfs-site.xml 에 구성된 Amazon SQS 대기열 이름입니다. 기본값은 **EMRFS-Inconsistency-<jobFlowId>** 입니다.	예

EMRFS CLI 명령을 단계로 제출

다음 예제에서는 AWS CLI 또는 API와를 활용하여 emrfs 명령을 단계로 실행하여 마스터 노드에서 emrfs 유틸리티command-runner.jar를 사용하는 방법을 보여줍니다. 이 예제에서는 AWS SDK for Python (Boto3) 를 사용하여 Amazon S3 버킷의 객체를 기본 EMRFS 메타데이터 테이블에 추가하는 단계를 클러스터에 추가합니다.

```
import boto3
from botocore.exceptions import ClientError


def add_emrfs_step(command, bucket_url, cluster_id, emr_client):
    """
    Add an EMRFS command as a job flow step to an existing cluster.

    :param command: The EMRFS command to run.
    :param bucket_url: The URL of a bucket that contains tracking metadata.
    :param cluster_id: The ID of the cluster to update.
    :param emr_client: The Boto3 Amazon EMR client object.
    :return: The ID of the added job flow step. Status can be tracked by calling
             the emr_client.describe_step() function.
    """
    job_flow_step = {
        "Name": "Example EMRFS Command Step",
        "ActionOnFailure": "CONTINUE",
        "HadoopJarStep": {
            "Jar": "command-runner.jar",
            "Args": ["/usr/bin/emrfs", command, bucket_url],
        },
    }

    try:
        response = emr_client.add_job_flow_steps(
            JobFlowId=cluster_id, Steps=[job_flow_step]
        )
```

```
        step_id = response["StepIds"][0]
        print(f"Added step {step_id} to cluster {cluster_id}.")
    except ClientError:
        print(f"Couldn't add a step to cluster {cluster_id}.")
        raise
    else:
        return step_id


def usage_demo():
    emr_client = boto3.client("emr")
    # Assumes the first waiting cluster has EMRFS enabled and has created metadata
    # with the default name of 'EmrFSMetadata'.
    cluster = emr_client.list_clusters(ClusterStates=["WAITING"])["Clusters"][0]
    add_emrfs_step(
        "sync", "s3://elasticmapreduce/samples/cloudfront", cluster["Id"], emr_client
    )


if __name__ == "__main__":
    usage_demo()
```

반환된 `step_id` 값을 사용하여 로그에서 작업 결과를 확인할 수 있습니다.

Amazon S3에서 EMRFS 데이터에 대한 액세스 권한 부여

기본적으로 EC2의 EMR 역할에 따라 Amazon S3에서 EMRFS 데이터를 액세스할 수 있는 권한이 결정됩니다. EMRFS를 통해 요청을 수행하는 사용자나 그룹에 관계없이 이 역할에 연결된 IAM 정책이 적용됩니다. 기본값은 `EMR_EC2_DefaultRole`입니다. 자세한 내용은 클러스터 EC2 인스턴스의 서비스 역할(EC2 인스턴스 프로파일)을 참조하세요.

Amazon EMR 릴리스 버전 5.10.0부터 보안 구성을 사용하여 EMRFS의 IAM 역할을 지정할 수 있습니다. 이 기능을 사용하면 사용자가 여러 명인 클러스터에서 Amazon S3에 대한 EMRFS 요청 권한을 사용자 지정할 수 있습니다. 사용자 및 그룹마다, 그리고 Amazon S3에서 접두사에 따라 Amazon S3 버킷 위치마다 다른 IAM 역할을 지정할 수 있습니다. EMRFS가 지정된 사용자, 그룹 또는 위치와 일치하는 Amazon S3에 대한 요청을 수행할 때 클러스터에서는 EC2의 EMR 역할 대신 사용자가 지정한 해당 역할을 사용합니다. 자세한 내용은 Amazon S3에 대한 EMRFS 요청의 IAM 역할 구성을 참조하세요.

또는 Amazon EMR 솔루션의 요구가 EMRFS의 IAM 역할이 제공하는 범위를 벗어난 경우 사용자 지정 보안 인증 제공업체 클래스를 정의할 수 있기 때문에 Amazon S3에서 EMRFS 데이터에 대한 액세스 권한을 사용자가 지정할 수 있습니다.

Amazon S3의 EMRFS 데이터에 대한 사용자 지정 보안 인증 제공업체 생성

사용자 지정 자격 증명 공급자를 생성하려면 AWSCredentialsProvider 및 하둡 구성 가능 클래스를 구현합니다.

이 접근 방식에 대한 자세한 설명은 빅 데이터 블로그의 EMRFS를 사용하여 다른 AWS 계정의 데이터를 안전하게 분석을 참조하세요 AWS . 블로그 게시물에는 IAM 역할 생성부터 클러스터 시작에 이르기까지 전체 프로세스를 안내하는 자습서가 포함되어 있습니다. 이 게시물은 사용자 지정 자격 증명 공급자 클래스를 구현하는 Java 코드 예제도 제공합니다.

기본 단계는 다음과 같습니다.

사용자 지정 자격 증명 공급자를 지정하려면

1. JAR 파일로 컴파일된 사용자 지정 자격 증명 공급자 클래스를 생성합니다.
2. 스크립트를 부트스트랩 작업으로 실행해 사용자 지정 자격 증명 공급자 JAR 파일을 클러스터의 마스터 노드에 있는 `/usr/share/aws/emr/emrfs/auxlib` 위치로 복사합니다. 부트스트랩 작업에 대한 자세한 내용은 (선택 사항)부트스트랩 작업을 생성하여 추가 소프트웨어 설치를 참조하세요.
3. `emrfs-site` 분류를 사용자 지정하여 JAR 파일에서 구현되는 클래스를 지정합니다. 애플리케이션을 사용자 지정하기 위해 구성 객체를 지정하는 방법에 대한 자세한 내용은 Amazon EMR 릴리스 안내서에서 애플리케이션 구성을 참조하세요.

   다음 예제는 일반적인 구성 파라미터를 사용하여 Hive 클러스터를 시작하는 `create-cluster` 명령을 보여줍니다.

   - Amazon S3의 *amzn-s3-demo-bucket*에 저장되는 스크립트인 *copy_jar_file.sh*를 실행하는 부트스트랩 작업.
   - JAR 파일에 정의된 사용자 지정 자격 증명 공급자를 `emrfs-site`로 지정하는 *MyCustomCredentialsProvider* 분류

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --applications Name=Hive \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket/
copy_jar_file.sh","Name":"Custom action"}]' \
--ec2-attributes '{"KeyName":"MyKeyPair","InstanceProfile":"EMR_EC2_DefaultRole",\
"SubnetId":"subnet-xxxxxxxx","EmrManagedSlaveSecurityGroup":"sg-xxxxxxxx",\
"EmrManagedMasterSecurityGroup":"sg-xxxxxxxx"}' \
--service-role EMR_DefaultRole_V2 --enable-debugging --release-label emr-7.9.0 \
--log-uri 's3n://amzn-s3-demo-bucket/' --name 'test-awscredentialsprovider-emrfs' \
--instance-type=m5.xlarge --instance-count 3  \
--configurations '[{"Classification":"emrfs-site",\
"Properties":
{"fs.s3.customAWSCredentialsProvider":"MyAWSCredentialsProviderWithUri"},\
"Configurations":[]}]'
```

기본 AWS Security Token Service 엔드포인트 관리

EMRFS는 AWS Security Token Service (STS)를 사용하여 AWS 리소스에 액세스하기 위해 임시 보안 자격 증명을 검색합니다. 이전 Amazon EMR 릴리스 버전은의 단일 글로벌 엔드포인트로 모든 AWS STS 요청을 보냅니다`https://sts.amazonaws.com`. Amazon EMR 릴리스 버전 5.31.0 및 6.1.0 이상에서는 리전 AWS STS 엔드포인트에 대신 요청합니다. 이렇게 하면 지연 시간이 줄어들고 세션 토큰 유효성이 개선됩니다. 엔드포인트에 대한 AWS STS 자세한 내용은 AWS Identity and Access Management 사용 설명서의 [AWS 리전 AWS STS 에서 관리를](#) 참조하세요.

Amazon EMR 릴리스 버전 5.31.0 및 6.1.0 이상을 사용하는 경우 기본 AWS STS 엔드포인트를 재정의할 수 있습니다. 이렇게 하려면 `emrfs-site` 구성에서 `fs.s3.sts.endpoint` 속성을 변경해야 합니다.

다음 AWS CLI 예제에서는 EMRFS에서 사용하는 기본 AWS STS 엔드포인트를 글로벌 엔드포인트로 설정합니다.

```
aws emr create-cluster --release-label <emr-5.33.0> --instance-type m5.xlarge \
--emrfs Args=[fs.s3.sts.endpoint=https://sts.amazonaws.com]
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

또는 다음 예제를 사용하여 JSON 구성 파일을 생성하고 `--configurations`의 `emr create-cluster` 인수를 사용하여 지정할 수 있습니다. `--configurations`, 사용에 대한 자세한 내용은 AWS CLI 명령 참조를 참조하세요.

```
[
  {
    "classification": "emrfs-site",
    "properties": {
      "fs.s3.sts.endpoint": "https://sts.amazonaws.com"
    }
  }
]
```

EMRFS 속성을 사용하여 Amazon S3 암호화 지정

Important

Amazon EMR 릴리스 버전 4.8.0부터 보안 구성을 사용하여 더 쉽게 더 많은 옵션을 통해 암호화 설정을 적용할 수 있습니다. 보안 구성을 사용하는 것이 좋습니다. 자세한 내용은 데이터 암호화 구성을 참조하세요. 이 단원에 설명되어 있는 콘솔 지침은 릴리스 버전 4.8.0 이하에서 사용이 가능합니다. AWS CLI 를 사용하여 클러스터 구성과 후속 버전의 보안 구성 모두에서 Amazon S3 암호화를 구성하는 경우 보안 구성이 클러스터 구성을 재정의합니다.

클러스터를 생성할 때 콘솔을 사용하거나 AWS CLI 또는 EMR SDK를 통해 `emrfs-site` 분류 속성을 사용하여 Amazon S3의 EMRFS 데이터에 대해 서버 측 암호화(SSE) 또는 클라이언트 측 암호화(CSE)를 지정할 수 있습니다. Amazon S3 SSE 및 CSE는 상호 배타적이므로 둘 중 하나를 선택할 수 있지만 둘 다 선택할 수는 없습니다.

AWS CLI 지침은 아래의 암호화 유형에 해당하는 섹션을 참조하세요.

를 사용하여 EMRFS 암호화 옵션을 지정하려면 AWS Management Console

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. 릴리스로 4.7.2 이전 버전을 선택합니다.

4. Software and Steps(소프트웨어 및 단계)에서 애플리케이션에 적절한 다른 옵션을 선택하고 다음을 선택합니다.
5. 하드웨어 및 General Cluster Settings(일반 클러스터 설정) 창에서 애플리케이션에 적절한 설정을 선택합니다.
6. 보안 창의 인증 및 암호화에서 사용하고 싶은 S3 Encryption (with EMRFS)(S3 암호화(EMRFS 사용)) 옵션을 선택합니다.

> **Note**
>
> Amazon EMR 릴리스 버전 4.4 이하를 사용할 때는 KMS 키 관리를 사용한 S3 서버 측 암호화(SSE-KMS)를 사용할 수 없습니다.

   - AWS Key Management를 사용하는 옵션을 선택할 경우 AWS KMS 키 ID를 선택합니다. 자세한 내용은 EMRFS 암호화 AWS KMS keys 에 사용 단원을 참조하십시오.
   - S3 client-side encryption with custom materials provider(사용자 지정 자료 공급자를 사용한 S3 클라이언트 측 암호화)를 선택한 경우에는 클래스 이름과 JAR location(JAR 위치)을 입력합니다. 자세한 내용은 Amazon S3 클라이언트 측 암호화 단원을 참조하십시오.
7. 애플리케이션에 적절한 다른 옵션을 선택하고 클러스터 생성을 선택합니다.

EMRFS 암호화 AWS KMS keys 에 사용

AWS KMS 암호화 키는 Amazon EMR 클러스터 인스턴스 및 EMRFS와 함께 사용되는 Amazon S3 버킷과 동일한 리전에서 생성해야 합니다. 지정하는 키가 클러스터를 구성하는 데 사용하는 계정이 아닌 다른 계정에 있는 경우 ARN을 사용하여 해당 키를 지정해야 합니다.

Amazon EC2 인스턴스 프로파일의 역할에는 사용자가 지정하는 CMK를 사용할 수 있는 권한이 있어야 합니다. Amazon EMR에 있는 인스턴스 프로파일의 기본 역할은 `EMR_EC2_DefaultRole`입니다. 인스턴스 프로파일에 다른 역할을 사용하거나 Amazon S3에 대한 EMRFS 요청에 IAM 역할을 사용하는 경우 각 역할이 키 사용자로 적절히 추가되도록 해야 합니다. 이렇게 하면 해당 역할에 KMS 키를 사용할 수 있는 권한이 부여됩니다. 자세한 내용은 AWS Key Management Service 개발자 안내서에서 Using Key Policies 및 Amazon S3에 대한 EMRFS 요청의 IAM 역할 구성을 참조하세요.

AWS Management Console 를 사용하여 지정된 KMS 키의 키 사용자 목록에 인스턴스 프로파일 또는 EC2 인스턴스 프로파일을 추가하거나 AWS CLI 또는 AWS SDK를 사용하여 적절한 키 정책을 연결할 수 있습니다.

Amazon EMR은 [대칭 KMS 키]만 지원합니다. [비대칭 KMS 키]를 사용하여 Amazon EMR 클러스터의 저장된 데이터를 암호화할 수 없습니다. KMS 키가 대칭인지 비대칭인지 확인하는 것과 관련된 도움말은 [Identifying symmetric and asymmetric KMS keys]를 참조하세요.

아래 절차에서는 AWS Management Console을 사용하여 기본 Amazon EMR 인스턴스 프로파일(`EMR_EC2_DefaultRole`)을 키 사용자로 추가하는 방법을 설명합니다. 이때 KMS 키를 이미 생성했다고 가정합니다. 새 KMS 키를 생성하려면 AWS Key Management Service 개발자 안내서에서 [Creating Keys]를 참조하세요.

Amazon EMR의 EC2 인스턴스 프로파일을 암호화 키 사용자 목록에 추가하는 방법

1. 에 로그인 AWS Management Console 하고 https://console.aws.amazon.com/kms AWS Key Management Service (AWS KMS) 콘솔을 엽니다.
2. 를 변경하려면 페이지 오른쪽 상단에 있는 리전 선택기를 AWS 리전사용합니다.
3. 수정할 KMS 키의 별칭을 선택합니다.
4. 키 세부 정보 페이지의 Key Users(키 사용자)에서 Add(추가)를 선택합니다.
5. Add key users(키 사용자 추가) 대화 상자에서 적절한 역할을 선택합니다. 기본 역할의 이름은 `EMR_EC2_DefaultRole`입니다.
6. 추가를 선택합니다.

Amazon S3 서버 측 암호화

Amazon S3 서버 측 암호화를 설정하면 Amazon S3에서 데이터를 디스크에 쓸 때 객체 수준에서 데이터를 암호화하고 데이터에 액세스할 때 데이터의 암호를 해독합니다. SSE에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 [서버 측 암호화를 사용하여 데이터 보호]를 참조하세요.

Amazon EMR에서 SSE를 지정할 때 다음 두 가지 키 관리 시스템 중에서 선택할 수 있습니다.

- SSE-S3 - Amazon S3에서 자동으로 키를 관리합니다.
- SSE-KMS - AWS KMS key 를 사용하여 Amazon EMR에 적합한 정책을 설정합니다. Amazon EMR의 키 요구 사항에 대한 자세한 내용은 [암호화에 사용을 참조 AWS KMS keys 하세요].

고객 제공 키를 사용하는 SSE(SSE-C)는 Amazon EMR에서 사용할 수 없습니다.

를 사용하여 SSE-S3가 활성화된 클러스터를 생성하려면 AWS CLI

- 다음 명령을 입력합니다.

```
aws emr create-cluster --release-label emr-4.7.2 or earlier \
--instance-count 3 --instance-type m5.xlarge --emrfs Encryption=ServerSide
```

emrfs-site 속성에서 fs.s3.enableServerSideEncryption 속성을 true로 설정하여 SSE-S3을 활성화할 수도 있습니다. 아래의 SSE-KMS 예제를 참조하여 키 ID에 대한 속성을 생략하십시오.

를 사용하여 SSE-KMS가 활성화된 클러스터를 생성하려면 AWS CLI

> **Note**
> SSE-KMS는 Amazon EMR 릴리스 버전 4.5.0 이상에서만 사용이 가능합니다.

- 다음 AWS CLI 명령을 입력하여 SSE-KMS로 클러스터를 생성합니다. 여기서 *keyID*는 *a4567b8-9900-12ab-1234-123a45678901* AWS KMS key과 같은 입니다.

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3 \
--instance-type m5.xlarge --use-default-roles \
--emrfs Encryption=ServerSide,Args=[fs.s3.serverSideEncryption.kms.keyId=keyId]
```

--또는--

emrfs-site 분류를 사용하여 다음 AWS CLI 명령을 입력하고 아래 예제와 유사한 내용이 포함된 구성 JSON 파일을 제공합니다myConfig.json.

```
aws emr create-cluster --release-label emr-4.7.2 or earlier --instance-count 3
 --instance-type m5.xlarge --applications Name=Hadoop --configurations file://
myConfig.json --use-default-roles
```

다음은 myConfig.json 콘텐츠 예제입니다.

```
[
  {
    "Classification":"emrfs-site",
    "Properties": {
```

```
        "fs.s3.enableServerSideEncryption": "true",
        "fs.s3.serverSideEncryption.kms.keyId":"a4567b8-9900-12ab-1234-123a45678901"
      }
    }
]
```

SSE-S3 및 SSE-KMS의 구성 속성

`emrfs-site` 구성 분류를 사용하여 이들 속성을 구성할 수 있습니다. SSE-KMS는 Amazon EMR 릴리스 버전 4.5.0 이상에서만 사용이 가능합니다.

속성	기본값	설명
`fs.s3.enableServerSideEncryption`	**false**	**true**로 설정할 경우 Amazon S3에 저장된 객체가 서버 측 암호화를 사용하여 암호화됩니다. 키를 지정하지 않으면 SSE-S3가 사용됩니다.
`fs.s3.serverSideEncryption.kms.keyId`	**n/a**	AWS KMS 키 ID 또는 ARN을 지정합니다. 키를 지정하면 SSE-KMS가 사용됩니다.

Amazon S3 클라이언트 측 암호화

Amazon S3 클라이언트 측 암호화를 사용하면 클러스터의 EMRFS 클라이언트에서 Amazon S3 암호화 및 암호 해독이 수행됩니다. 객체는 Amazon S3에 업로드되기 전에 암호화되고 다운로드된 후 암호 해독됩니다. 지정하는 공급자는 클라이언트가 사용하는 암호화 키를 제공합니다. 클라이언트는 AWS KMS 에서 제공하는 키(CSE-KMS) 또는 클라이언트 측 루트 키(CSE-C)를 제공하는 사용자 지정 Java 클래스를 사용할 수 있습니다. 암호화 세부 사항은 지정된 공급자 및 암호 해독되거나 암호화되는 객체의 메타데이터에 따라 CSE-KMS 및 CSE-C 간에 약간 다릅니다. 이러한 차이에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서의 클라이언트 측 암호화를 사용하여 데이터 보호를 참조하세요.

Note

Amazon S3 CSE는 Amazon S3와 교환하는 EMRFS 데이터만 암호화하며, 클러스터 인스턴스 볼륨에 있는 모든 데이터를 암호화하지는 않습니다. 뿐만 아니라, Hue에서 EMRFS가 사용되지 않으므로 Hue S3 파일 검색기를 사용하여 Amazon S3에 작성된 객체는 암호화되지 않습니다.

를 사용하여 Amazon S3에서 EMRFS 데이터에 대한 CSE-KMS를 지정하려면 AWS CLI

- 다음 명령을 입력하고 *MyKMSKeyID*를 사용할 KMS 키의 키 ID나 ARN으로 바꿉니다.

```
aws emr create-cluster --release-label emr-4.7.2 or earlier
--emrfs Encryption=ClientSide,ProviderType=KMS,KMSKeyId=MyKMSKeyId
```

사용자 지정 키 공급자 생성

사용자 지정 키 제공업체를 생성할 때 사용하는 암호화 유형에 따라 애플리케이션은 서로 다른 EncryptionMaterialsProvider 인터페이스도 구현해야 합니다. 두 인터페이스 모두 AWS SDK for Java 버전 1.11.0 이상에서 사용할 수 있습니다.

- Amazon S3 암호화를 구현하려면 com.amazonaws.services.s3.model.EncryptionMaterialsProvider 인터페이스를 사용합니다.
- 로컬 디스크 암호화를 구현하려면 com.amazonaws.services.elasticmapreduce.spi.security.EncryptionMaterialsProvider 인터페이스를 사용합니다.

이 구현을 위해 암호화 구성 요소를 제공하도록 임의의 모든 전략을 사용할 수 있습니다. 예를 들어 정적 암호화 구성 요소를 제공하거나 더 복잡한 키 관리 시스템과 통합하도록 선택할 수 있습니다.

Amazon S3 암호화를 사용하는 경우 사용자 지정 암호화 구성 요소에 대해 암호화 알고리즘 AES/GCM/NoPadding을 사용해야 합니다.

로컬 디스크 암호화를 사용하는 경우 사용자 지정 암호화 구성 요소에 사용할 암호화 알고리즘은 EMR 릴리스에 따라 다릅니다. Amazon EMR 7.0.0 이하의 경우 AES/GCM/NoPadding을 사용해야 합니다. Amazon EMR 7.1.0 이상의 경우 AES를 사용해야 합니다.

EncryptionMaterialsProvider 클래스는 암호화 컨텍스트별로 암호화 자료를 가져옵니다. Amazon EMR은 직접 호출자가 반환할 올바른 암호화 자료를 결정하는 데 도움이 되도록 런타임에 암호화 컨텍스트 정보를 채웁니다.

Example 예: EMRFS에서 Amazon S3 암호화를 위해 사용자 지정 키 공급자 사용

Amazon EMR이 EncryptionMaterialsProvider 클래스에서 암호화 구성 요소를 가져와서 암호화를 수행할 때 EMRFS는 선택적으로 materialsDescription 인수를 객체의 Amazon S3 URI 및 클러스터의 JobFlowId라는 두 개의 필드로 채웁니다. 이러한 필드는 EncryptionMaterialsProvider 클래스에서 암호화 구성 요소를 선택적으로 반환하기 위해 사용할 수 있습니다.

예를 들어, 공급자는 Amazon S3 URI 접두사마다 다른 키를 반환할 수 있습니다. Amazon S3 객체에 저장되는 것은 EMRFS에서 생성되어 공급자에게 전달되는 materialsDescription 값이 아니라 반환되는 암호화 구성 요소에 대한 설명입니다. Amazon S3 객체를 암호 해독하는 동안, 암호화 구성 요소 설명이 EncryptionMaterialsProvider 클래스에 전달되므로, 이 경우에도 일치하는 키를 반환하여 객체를 암호 해독할 수 있습니다.

EncryptionMaterialsProvider 참조 구현은 아래에 제공됩니다. 또 다른 사용자 지정 공급자인 EMRFSRSAEncryptionMaterialsProvider는 GitHub에서 사용할 수 있습니다.

```
import com.amazonaws.services.s3.model.EncryptionMaterials;
import com.amazonaws.services.s3.model.EncryptionMaterialsProvider;
import com.amazonaws.services.s3.model.KMSEncryptionMaterials;
import org.apache.hadoop.conf.Configurable;
import org.apache.hadoop.conf.Configuration;

import java.util.Map;

/**
 * Provides KMSEncryptionMaterials according to Configuration
 */
public class MyEncryptionMaterialsProviders implements EncryptionMaterialsProvider,
 Configurable{
  private Configuration conf;
  private String kmsKeyId;
  private EncryptionMaterials encryptionMaterials;

  private void init() {
    this.kmsKeyId = conf.get("my.kms.key.id");
    this.encryptionMaterials = new KMSEncryptionMaterials(kmsKeyId);
  }
```

```
  @Override
  public void setConf(Configuration conf) {
    this.conf = conf;
    init();
  }

  @Override
  public Configuration getConf() {
    return this.conf;
  }

  @Override
  public void refresh() {

  }

  @Override
  public EncryptionMaterials getEncryptionMaterials(Map<String, String>
 materialsDescription) {
    return this.encryptionMaterials;
  }

  @Override
  public EncryptionMaterials getEncryptionMaterials() {
    return this.encryptionMaterials;
  }
}
```

를 사용하여 사용자 지정 재료 공급자 지정 AWS CLI

AWS CLI를 사용하려면 `Encryption`, `ProviderType`, `CustomProviderClass` 및`CustomProviderLocation` 인수를 `emrfs` 옵션으로 전달합니다.

```
aws emr create-cluster --instance-type m5.xlarge --release-label emr-4.7.2 or earlier
 --emrfs Encryption=ClientSide,ProviderType=Custom,CustomProviderLocation=s3://amzn-s3-
demo-bucket/myfolder/provider.jar,CustomProviderClass=classname
```

`Encryption`을 `ClientSide`로 설정하면 클라이언트 측 암호화가 활성화됩니다. `CustomProviderClass`는 `EncryptionMaterialsProvider`객체의 이름이고 `CustomProviderLocation`은 Amazon EMR이 `CustomProviderClass`를 클러스터의 각 노드로 복사하여 클래스 경로에 저장하는 로컬 또는 Amazon S3 위치입니다.

SDK를 사용하여 사용자 지정 자격 공급자 지정

SDK를 사용하려면 먼저 `fs.s3.cse.encryptionMaterialsProvider.uri` 속성을 설정하여 Amazon S3에 저장하는 사용자 지정 `EncryptionMaterialsProvider` 클래스를 클러스터의 각 노드로 다운로드합니다. 그런 다음 사용자 지정 공급자의 올바른 위치 및 CSE 활성화와 함께 `emrfs-site.xml` 파일에서 이 동작을 구성합니다.

예를 들어 RunJobFlowRequest를 AWS SDK for Java 사용하는에서 코드는 다음과 같을 수 있습니다.

```
<snip>
  Map<String,String> emrfsProperties = new HashMap<String,String>();
      emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider.uri","s3://amzn-s3-
demo-bucket/MyCustomEncryptionMaterialsProvider.jar");
      emrfsProperties.put("fs.s3.cse.enabled","true");
      emrfsProperties.put("fs.s3.consistent","true");

 emrfsProperties.put("fs.s3.cse.encryptionMaterialsProvider","full.class.name.of.EncryptionMate

  Configuration myEmrfsConfig = new Configuration()
      .withClassification("emrfs-site")
      .withProperties(emrfsProperties);

  RunJobFlowRequest request = new RunJobFlowRequest()
   .withName("Custom EncryptionMaterialsProvider")
   .withReleaseLabel("emr-7.9.0")
   .withApplications(myApp)
   .withConfigurations(myEmrfsConfig)
   .withServiceRole("EMR_DefaultRole_V2")
   .withJobFlowRole("EMR_EC2_DefaultRole")
   .withLogUri("s3://myLogUri/")
   .withInstances(new JobFlowInstancesConfig()
    .withEc2KeyName("myEc2Key")
    .withInstanceCount(2)
    .withKeepJobFlowAliveWhenNoSteps(true)
    .withMasterInstanceType("m5.xlarge")
    .withSlaveInstanceType("m5.xlarge")
   );

  RunJobFlowResult result = emr.runJobFlow(request);
</snip>
```

사용자 지정 EncryptionMaterialsProvider(인수 포함)

공급자에게 직접 인수를 전달해야 할 수 있습니다. 이를 위해 사용자 지정 인수를 속성으로 정의한 상태에서 `emrfs-site` 구성 분류를 사용할 수 있습니다. 구성의 예제는 아래와 같습니다(파일 `myConfig.json`로 저장).

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "myProvider.arg1":"value1",
     "myProvider.arg2":"value2"
      }
    }
 ]
```

의 `create-cluster` 명령을 사용하여 옵션을 사용하여 아래와 같이 파일을 `--configurations` 지정할 AWS CLI수 있습니다.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge
 --instance-count 2 --configurations file://myConfig.json --emrfs
 Encryption=ClientSide,CustomProviderLocation=s3://amzn-s3-demo-bucket/myfolder/
myprovider.jar,CustomProviderClass=classname
```

EMRFS S3EC V2 지원 구성

S3 Java SDK 릴리스(1.11.837 이상)에서 다양한 보안 개선 사항과 함께 암호화 클라이언트 버전 2(S3EC V2)를 지원합니다. 자세한 내용은 S3 블로그 게시물 Updates to the Amazon S3 encryption client를 참조하세요. 또한 AWS SDK for Java 개발자 안내서의 Amazon S3 암호화 클라이언트 마이그레이션을 참조하세요.

암호화 클라이언트 V1은 이전 버전과의 호환성을 위해 SDK에서 계속 사용할 수 있습니다. 기본적으로 EMRFS는 CSE가 활성화된 경우 S3EC V1을 사용하여 S3 객체를 암호화하고 해독합니다.

릴리스 버전이 emr-5.31.0 이전(emr-5.30.1 이하, emr-6.1.0 이하)인 EMR 클러스터에서는 S3EC V2로 암호화된 S3 객체를 EMRFS로 해독할 수 없습니다.

Example S3EC V2를 사용하도록 EMRFS 구성

S3EC V2를 사용하도록 EMRFS를 구성하려면 다음 구성을 추가합니다.

```
{
  "Classification": "emrfs-site",
  "Properties": {
    "fs.s3.cse.encryptionV2.enabled": "true"
  }
}
```

Amazon S3 클라이언트 측 암호화에 대한 **emrfs-site.xml** 속성

속성	기본값	설명
fs.s3.cse.enabled	**false**	**true**로 설정할 경우 Amazon S3에 저장된 EMRFS 객체가 클라이언트 측 암호화를 사용하여 암호화됩니다.
fs.s3.cse.encryptionV2.enabled	**false**	true로 설정할 경우 EMRFS는 S3 암호화 클라이언트 버전 2를 사용하여 S3의 객체를 암호화하고 해독합니다. Amazon EMR 버전 5.31.0 이상에서 사용할 수 있습니다.
fs.s3.cse.encryptionMaterialsProvider.uri	**N/A**	사용자 지정 암호화 자료를 사용할 때 이렇게 설정합니다. EncryptionMaterialsProvider 가 있는 JAR이 위치한 Amazon S3 URI. 이 URI를 제공할 경우 Amazon EMR에서 클러스터의 모든 노드에 JAR을 자동으로 다운로드합니다.
fs.s3.cse.encryptionMaterialsProvider	**N/A**	클라이언트 측 암호화에 사용되는 EncryptionMaterialsProvider 클래스 경로입니다. CSE-KMS를 사용하

속성	기본값	설명
		여 `com.amazon.ws.emr.hadoop.fs.cse.KMSEncryptionMaterialsProvider` 를 지정합니다.
`fs.s3.cse.materialsDescription.enabled`	**`false`**	`true`로 설정할 경우 암호화된 객체의 materialsDescription이 객체의 Amazon S3 URI와 JobFlowId로 채워집니다. 사용자 지정 암호화 자료를 사용할 때 `true`로 설정합니다.
`fs.s3.cse.kms.keyId`	**`N/A`**	CSE-KMS를 사용할 때 적용됩니다. 암호화에 사용되는 KMS 키의 KeyId, ARN 또는 별칭에 대한 값.
`fs.s3.cse.cryptoStorageMode`	**`ObjectMetadata`**	Amazon S3 스토리지 모드. 기본적으로 암호화 정보의 설명은 객체 메타데이터에 저장됩니다. 지침 파일에 설명을 저장할 수도 있습니다. 유효한 값은 ObjectMetadata 및 InstructionFile입니다. 자세한 내용은 [AWS SDK for Java 및 Amazon S3를 사용한 클라이언트 측 데이터 암호화를](#) 참조하세요.

S3A 파일 시스템

이 섹션에서는 S3A 파일 시스템을 사용할 때 Amazon Elastic Map Reduce(EMR)에서 실행되는 Spark에 대한 프로토콜을 다룹니다.

S3A MagicV2 커미터

EMR-6.15.0 릴리스를 통해 Amazon EMR은 MagicV2 커미터라고 하는 새로운 S3A 커미터 유형을 도입합니다. 이 기능에 대한 자세한 내용은 관련 설명서 섹션을 참조하세요.

MagicV2 커미터는 S3A 파일 시스템을 통해 Amazon S3에 대한 파일 쓰기를 최적화하도록 특별히 설계된 오픈 소스 MagicCommitter의 향상된 구현을 나타냅니다. 이전 버전과 마찬가지로 Amazon S3의 멀티파트 업로드 기능을 활용하여 일반적으로 작업 및 작업 커밋 단계와 관련된 기존 목록을 제거하고 작업의 이름을 변경합니다.

원래 MagicCommitter에 비해 MagicV2 커미터는 작업 커밋 단계가 아닌 작업 커밋 단계에서 작업의 출력 위치에 파일을 기록하여 우수한 성능을 보여줍니다. 이 접근 방식을 사용하면 분산 파일 쓰기가 가능하고 Amazon S3에서 임시 커밋 메타데이터 스토리지가 필요하지 않으므로 비용 효율성이 향상됩니다. 또한 MagicV2 커미터는 커밋 프로세스 중에 여러 스레드에서 파일 경로를 덮어쓸 수 있으므로 유연성이 향상됩니다.

MagicV2 커미터 활성화

MagicV2 커미터를 활성화하려면 작업 구성에 다음 구성을 전달하거나 코어 사이트 구성을 사용하여 속성을 설정합니다. 자세한 내용은 애플리케이션 구성을 참조하세요.

```
mapreduce.outputcommitter.factory.scheme.s3a=org.apache.hadoop.fs.s3a.commit.S3ACommitterFactor
fs.s3a.committer.magic.enabled=true
fs.s3a.committer.name=magicv2
fs.s3a.committer.magic.track.commits.in.memory.enabled=true
```

새 파일을 커밋하거나 쓰기 전에 기존 디렉터리를 덮어써야 하는 워크로드의 경우 앞서 언급한 구성과 함께 다음과 같은 추가 구성이 필요합니다.

```
fs.s3a.committer.magic.overwrite.and.commit=true
fs.s3a.committer.magic.delete.directory.threads=thread size
```

`threads` 구성의 기본값은 입니다`20`. 그러나 더 나은 성능을 위해 덮어쓸 디렉터리가 많은 경우이 파라미터를 조정해야 합니다. 이는 EMR-7.2.0 이상에서만 사용할 수 있습니다.

고려 사항

- 작업이 실행되고 Amazon S3에 데이터를 쓰는 동안 Java 가상 머신(JVM)이 충돌하거나 종료되면 불완전한 멀티파트 업로드가 뒤쳐질 가능성이 더 높습니다. 따라서 MagicV2 커미터를 사용할 때는 실패한 멀티파트 업로드를 관리하는 모범 사례를 따라야 합니다. 자세한 내용은 Amazon EMR 관리 안내서의 Amazon S3 버킷 작업 모범 사례 섹션을 참조하세요.
- 작업이 실패하더라도 성공한 작업에 의해 커밋된 모든 파일은 대상 경로에 계속 표시됩니다. 이러한 경우 사용자는 동일한 대상 경로에서 작업을 다시 실행하기 전에 커밋된 파일을 수동으로 정리해야 합니다.
- MagicV2 커미터는 작업이 커밋되거나 중단될 때까지 작업 시도로 작성된 각 파일에 대해 소량의 메모리를 사용합니다. 대부분의 작업에서 사용된 메모리 양은 무시해도 됩니다. 그러나 단일 실행기 프로세스가 많은 수의 작업을 동시에 처리하는 경우 메모리 압력이 높아지고 컨테이너 또는 실행기 메모리 부족(OOM)이 발생할 수 있습니다. 컨테이너 또는 실행기 메모리를 늘리면이 문제가 해결됩니다.

Amazon S3 S3A 클라이언트 측 암호화

Amazon Elastic Map Reduce(EMR) 릴리스 버전 7.6.0부터 S3A 파일 시스템 커넥터는 이제 Amazon S3 클라이언트 측 암호화를 지원합니다. 즉, Amazon S3 데이터의 암호화 및 복호화는 컴퓨팅 클러스터의 S3A 클라이언트 내에서 직접 이루어집니다. 이 기능을 사용하면 파일이 Amazon S3에 업로드되기 전에 자동으로 암호화되고 다운로드 시 해독됩니다. 암호화 방법론 및 구현에 대한 포괄적인 세부 정보는 Amazon Simple Storage Service 사용 설명서의 [클라이언트 측 암호화를 사용하여 데이터 보호를](#) 참조할 수 있습니다.

Amazon EMR에서 S3A를 사용하여 클라이언트 측 암호화(CSE)를 활성화하는 경우 두 가지 키 관리 시스템 옵션이 있습니다.

- CSE-KMS -이 접근 방식은 Amazon EMR용으로 특별히 설계된 정책으로 구성된 AWS Key Management Service (KMS) 키를 사용합니다. 키 요구 사항에 대한 자세한 내용은 [암호화에 AWS KMS 키 사용](#) 설명서를 참조하세요.
- CSE-CUSTOM -이 방법을 사용하면 데이터 암호화 및 복호화를 담당하는 클라이언트 측 루트 키를 제공하는 사용자 지정 Java 클래스를 통합할 수 있습니다.

> **Note**
>
> EMR의 S3A 클라이언트 측 암호화는 본질적으로 EMRFS 클라이언트 측 암호화와 호환되므로 EMRFS CSE를 사용하여 암호화된 객체는 S3A CSE를 통해 읽을 수 있습니다.

주제

- [CSE-KMS 설정](#)
- [CSE-사용자 지정 설정](#)
- [Amazon S3 S3A 클라이언트 측 암호화 속성](#)

CSE-KMS 설정

두 가지 기본 범위에서 AWS KMS (CSE-KMS)를 사용하여 클라이언트 측 암호화를 활성화할 수 있습니다.

- 첫 번째는 클러스터 전체 구성입니다.

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-KMS",
       "fs.s3a.encryption.key":"${KMS_KEY_ID}",
    }
  }
]
```

> **Note**
>
> AWS KMS 키 리전이 S3 버킷/EMR 리전과 다른 경우 추가 구성을 설정해야 합니다`fs.s3a.encryption.cse.kms.region=${KMS_REGION}`.

- 두 번째는 작업 또는 애플리케이션별 구성입니다. CSE-KMS는 다음과 같이 특정 Spark 애플리케이션에 대해 설정할 수 있습니다.

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-KMS --conf
 spark.hadoop.fs.s3a.encryption.key=${KMS_KEY_ID}
```

CSE-사용자 지정 설정

CSE-CUSTOM을 사용하려면 키링 인터페이스를 구현하여 사용자 지정 키 공급자를 생성해야 합니다. 다음은 샘플 구현입니다.

```
public class CustomKeyring implements Keyring {
  public CustomKeyring()  {
    // custom code
  }

  @Override
  public EncryptionMaterials onEncrypt(EncryptionMaterials encryptionMaterials) {
    // custom code
  }

  @Override
  public DecryptionMaterials onDecrypt(DecryptionMaterials decryptionMaterials,
      ListEncryptedDataKey list) {
```

```
    // custom code
  }
```

두 가지 기본 범위에서 클라이언트 측 암호화 사용자 지정 키(CSE-CUSTOM)를 활성화할 수 있습니다.

- 첫 번째 범위는 클러스터 전체 구성입니다.

```
[
  {
    "Classification":"core-site",
    "Properties": {
       "fs.s3a.encryption.algorithm": "CSE-CUSTOM",
       "fs.s3a.cse.customKeyringProvider.uri":"S3 path of custom jar",
       "fs.s3a.encryption.cse.custom.keyring.class.name":"fully qualified class name"
    }
  }
]
```

- 두 번째는 작업 또는 애플리케이션별 구성입니다. CSE-CUSTOM은 다음과 같이 특정 Spark 애플리케이션에 대해 설정할 수 있습니다.

```
spark-submit --conf spark.hadoop.fs.s3a.encryption.algorithm=CSE-CUSTOM --conf
 spark.hadoop.fs.s3a.encryption.cse.custom.keyring.class.name=fully qualified class
 name
```

> **Note**
>
> 암호화/복호화 키를 생성하는 데 필요한 사용자 지정 jar이 클래스 경로에 있는지 확인합니다.

Amazon S3 S3A 클라이언트 측 암호화 속성

S3A를 사용하여 클라이언트 측 암호화를 구성하려면 core-site.xml 설정에서 설정해야 하는 몇 가지 구성 속성이 있습니다. 사용자 지정 구성 설정에 대한 자세한 내용은 애플리케이션 구성을 참조하세요.

속성	기본값	설명
fs.s3a.encryption.algorithm	N/A	CSE-KMS 또는 로 설정하면 Amazon S3에 저장된 CSE-

속성	기본값	설명
		CUSTOM 객체는 클라이언트 측 암호화를 사용하여 암호화됩니다.
fs.s3a.encryption.key	N/A	를 사용할 때 적용됩니다CSE-KMS. 암호화에 사용되는 KMS 키의 KeyId, ARN 또는 별칭에 대한 값.
fs.s3a.encryption.cse.kms.region	N/A	를 사용할 때 적용됩니다CSE-KMS. AWS KMS 키가 생성되는 리전입니다. 기본적으로 KMS 리전은 S3 버킷/EMR 클러스터 리전과 동일한 값으로 설정됩니다.
fs.s3a.encryption.cse.custom.keyring.class.name	N/A	를 사용할 때 적용됩니다CSE-KMS. 사용자 지정 키 공급자의 정규화된 클래스 이름입니다.
fs.s3a.cse.customKeyringProvider.uri	N/A	를 사용할 때 적용됩니다CSE-CUSTOM . 키링의 사용자 지정 구현이 포함된 JAR이 있는 Amazon S3 URI입니다. 이 URI를 제공할 경우 Amazon EMR에서 클러스터의 모든 노드에 JAR을 자동으로 다운로드합니다.
fs.s3a.encryption.cse.v1.compatibility.enabled	'true'	이를 통해 EMRFS와 함께 사용되는 클라이언트와 같은 이전 SDK 클라이언트와의 이전 버전과의 호환성을 제공합니다. 이러한 종속성이 없는 경우 더 나은 성능을 위해이 기능을 끕니다.

Amazon CloudWatch 에이전트

Amazon EMR의 Amazon CloudWatch 에이전트는 EMR 클러스터의 Amazon EC2 인스턴스를 모니터링할 수 있는 도구입니다. CloudWatch에서 CloudWatch 에이전트로 수집한 지표를 저장하고 볼 수 있습니다. CloudWatch 에이전트에 대한 자세한 내용은 Amazon CloudWatch 사용 설명서를 참조하세요.

> **Note**
>
> 다른를 사용하여 Amazon CloudWatch 에이전트 지표 AWS 서비스 를 게시, 쿼리 또는 보는 경우 추가 요금이 발생합니다. 요금에 대한 자세한 내용은 다음 페이지를 참조하세요.
>
> - Amazon CloudWatch 요금
> - Amazon Managed Service for Prometheus 요금
> - Amazon Managed Grafana 요금
>
> 이러한 지표는 Amazon EMR이 CloudWatch의 `AWS/ElasticMapReduce` 네임스페이스에서 제공하는 무료 지표와는 별개입니다. CloudWatch 에이전트가 게시하지 않는 지표에 대한 자세한 내용은 CloudWatch를 사용하여 Amazon EMR 지표 모니터링을 참조하세요.

Amazon EMR 7.0 이상을 사용하면 클러스터에서 사용자 지정 버전의 Amazon CloudWatch 에이전트를 설치하여 EMR 클러스터에서 지표를 수집할 수 있습니다. CloudWatchAgent는 EMR 7.6 이상용 런타임 역할 클러스터에서 지원됩니다. 에이전트는 클러스터의 모든 노드에서 지표를 수집하고 프라이머리 노드에서 지표를 모아 해당 노드의 지표를 클라우드에 게시합니다.

콘솔에서 새 클러스터를 생성하거나 `create-cluster` API를 사용할 때 에이전트를 설치할 수 있습니다. 자세한 내용은 Amazon CloudWatch 에이전트를 사용하는 EMR 클러스터를 생성합니다. 단원을 참조하십시오. 기본적으로 릴리스 7.x 시리즈를 실행하는 Amazon EMR 클러스터는 60초 간격으로 34개의 시스템 수준 지표를 CloudWatch에 게시하지만 다른 지표를 게시하도록 에이전트를 구성할 수 있습니다. 또 다른 옵션은 Amazon Managed Service for Prometheus에 지표를 게시하는 것입니다. 이때 게시할 지표도 선택할 수 있습니다. 여러 사용 사례와 설정에 대해 CloudWatch 콘솔, Amazon Managed Grafana 또는 CloudWatch용 API나 Amazon Managed Service for Prometheus를 통해 지표를 보고 쿼리하도록 에이전트를 구성할 수 있습니다. 이러한 AWS 서비스 에서는 저장하고 쿼리하는 데 사용할 때 요금을 청구합니다.

Amazon EMR의 CloudWatch 에이전트는 Amazon EMR 클러스터에서 다음 서비스에 대한 시스템 지표와 JMX 지표를 게시할 수 있습니다.

- Hadoop DataNode
- Hadoop NameNode
- Yarn NodeManager
- Yarn ResourceManager
- HBase Master - Amazon EMR 7.1 이상만 해당
- HBase RegionServer - Amazon EMR 7.1 이상만 해당
- HBase ThriftServer – Amazon EMR 7.1 이상만 해당

사용 가능한 지표와 Amazon EMR에서 CloudWatch 에이전트를 구성하는 방법에 대한 자세한 내용은 Amazon EMR을 사용하도록 CloudWatch 에이전트 구성을 참조하세요.

다음 표에는 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 AmazonCloudWatchAgent 버전과 Amazon EMR이 AmazonCloudWatchAgent와 함께 설치하는 구성 요소가 나열되어 있습니다.

이 릴리스에서 AmazonCloudWatchAgent와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 AmazonCloudWatchAgent 버전 정보

Amazon EMR 릴리스 레이블	AmazonCloudWatchAgent 버전	AmazonCloudWatchAgent를 사용하여 설치한 구성 요소
emr-7.9.0	AmazonCloudWatchAgent 1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

주제

Amazon CloudWatch 에이전트를 사용하는 EMR 클러스터를 생성합니다.

이 섹션의 절차에서는 AWS Management Console 및에서 Amazon CloudWatch 에이전트를 사용하여 Amazon EMR에서 클러스터를 생성하는 단계를 설명합니다 AWS CLI.

주제

CloudWatch 에이전트에 대한 필수 IAM 권한

CloudWatch 에이전트에는 Amazon EMR용 Amazon EC2 인스턴스 프로파일에 AWS Identity and Access Management (IAM) `cloudwatch:PutMetricData` 권한이 필요합니다. Amazon EMR 기본 역할에는 이미 이 권한이 있습니다. 를 사용하여에서 기본 역할을 생성할 수 AWS CLI 있습니다`aws emr create-default-roles`. 자세한 내용은 Amazon EMR 관리 안내서에서 [클러스터 EC2 인스턴스의 서비스 역할(EC2 인스턴스 프로파일)]을 참조하세요.

다음 예시 IAM 정책에는 `cloudwatch:PutMetricData` 권한이 포함되어 있습니다.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
```

```
            "Action": "cloudwatch:PutMetricData",
            "Resource": "*"
        }
    ]
}
```

필수 CloudWatch 에이전트 엔드포인트

프라이빗 서브넷의 EMR 클러스터에 대한 지표를 CloudWatch에 게시하려면 CloudWatch 에이전트 엔드포인트를 생성하고 프라이빗 서브넷이 있는 VPC와 연결합니다.

각의 CloudWatch 엔드포인트에 대한 자세한 AWS 리전내용은 AWS 일반 참조 안내서의 Amazon CloudWatch 엔드포인트 및 할당량을 참조하세요.

EMR 클러스터 생성

CloudWatch 에이전트와 함께 사용하는 AWS CLI 데 필요한 권한과 엔드포인트를 설정한 후에는 AWS Management Console 또는를 사용하여 에이전트가 설치된 새 클러스터를 생성합니다.

Console

콘솔에서 Amazon CloudWatch 에이전트를 사용하여 클러스터를 만드는 방법

1. Amazon EMR 콘솔로 이동합니다.
2. 클러스터 생성을 선택합니다.
3. 이름 및 애플리케이션에서 Amazon EMR 릴리스 7.0.0 이상을 선택합니다.
4. 애플리케이션 번들에서 클러스터에 설치할 번들 또는 앱을 선택하고 선택 항목에 CloudWatch 에이전트를 포함합니다.
5. 사용 사례 요구 사항에 맞게 클러스터를 생성하세요.

AWS CLI

에서에 대한 `--applications` 파라미터를 사용하여 클러스터에 Amazon CloudWatch 에이전트를 추가할 AWS CLI수 있습니다`create-cluster`.

에서 Amazon CloudWatch 에이전트를 사용하여 클러스터를 생성하려면 AWS CLI

- 클러스터를 생성할 때 Amazon CloudWatch 에이전트를 포함하려면 다음과 비슷한 명령을 사용합니다. *`myKey`*를 EC2 키 페어로 바꿉니다.

```
aws emr create-cluster --name "Spark cluster with CloudWatch agent" \
--release-label emr-7.0.0 \
--applications Name=Spark Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--instance-count 3 --use-default-roles
```

에서 Amazon EMR을 사용하는 방법에 대한 자세한 AWS CLI내용은 AWS CLI 명령 참조를 참조하세요.

Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표

Amazon EMR에 Amazon CloudWatch 에이전트를 설치하면 에이전트를 다르게 구성하지 않는 한 기본 구성에 따라 클러스터의 모든 인스턴스에 대한 다음 시스템 지표가 게시됩니다. 각 지표의 정의는 Amazon CloudWatch 사용 설명서의 CloudWatch 에이전트가 수집하는 지표를 참조하세요.

CPU

CPU 지표

- cpu_usage_active
- cpu_usage_guest
- cpu_usage_guest_nice
- cpu_usage_idle
- cpu_usage_iowait
- cpu_usage_irq
- cpu_usage_nice
- cpu_usage_softirq
- cpu_usage_steal
- cpu_usage_system
- cpu_usage_user

Disk

디스크 지표

- disk_free
- disk_total
- disk_used
- disk_used_percent

Memory

메모리 지표

- mem_active
- mem_available
- mem_available_percent
- mem_free
- mem_inactive
- mem_total
- mem_used
- mem_used_percent
- mem_buffered
- mem_cached

Network IO

네트워크 IO 지표

- net_bytes_recv
- net_bytes_sent
- net_packets_recv
- net_packets_sent

Process

프로세스 지표

- processes_running
- processes_total

Swap

스왑 지표

- swap_free
- swap_used
- swap_used_percent

기본적으로 에이전트는 사용자 지정 네임스페이스 CWAgent 및 스키마 cluster.id, instance.id, node.type, service.name 아래에서 CloudWatch에 모든 지표를 게시합니다. 다음 단계를 사용하여 CloudWatch 에이전트가 Amazon EMR에 게시하는 이러한 기본 지표를 확인하세요.

1. Amazon CloudWatch 콘솔로 이동합니다.
2. 지표 탭을 선택한 다음 모든 지표를 선택합니다.
3. 사용자 지정 네임스페이스에서 **CWAgent**를 선택합니다. 그런 다음, **cluster.id**, **instance.id**, **node.type**, **service.name** 스키마를 선택합니다.
4. 사용 사례에 따라 CloudWatch 인터페이스에서 지표를 계속 쿼리하세요.

Amazon EMR은 각 지표에 ["cluster.id", "instance.id', "node.type", "service.name"] 레이블을 연결합니다. 이러한 레이블과 관련하여 다음 사항을 고려하세요.

- cluster.id 레이블 값은 지표를 생성한 EMR 클러스터의 ID입니다. cluster.id 레이블 값의 예는 j-123456789ABC입니다.
- instance.id 레이블 값은 지표를 생성한 EMR 클러스터의 인스턴스 ID입니다. instance.id 레이블 값의 예는 i-01bcf5f140f3355777입니다.
- node.type 레이블 값은 지표를 생성한 클러스터의 인스턴스에 대한 노드 유형을 나타냅니다. 이 값은 master, core 또는 task일 수 있습니다.
- service.name 레이블 값은 지표를 생성한 서비스의 이름입니다. 기본 service.name 시스템 지표의 기본값은 system입니다.

> **Note**
>
> Amazon EMR 릴리스 7.0.0은 약간 다른 지표 레이블 스키마를 사용합니다. `cluster.id`는 `jobflow.id`이고, `node.type`은 사용할 수 없습니다. `instance.id` 및 `service.name`은 사용할 수 있습니다.

Amazon EMR을 사용하도록 CloudWatch 에이전트 구성

Amazon EMR 7.0.0 이상에는 Amazon CloudWatch 에이전트가 포함됩니다. 다음 섹션에서는 Amazon CloudWatch 에이전트를 사용하는 EMR 클러스터를 생성합니다.에 설명된 대로 EMR 클러스터에서 에이전트를 구성하는 방법을 설명합니다.

주제

- Amazon EMR 7.1.0용 CloudWatch 에이전트 구성
- Amazon EMR 7.0.0용 CloudWatch 에이전트 구성

Amazon EMR 7.1.0용 CloudWatch 에이전트 구성

Amazon EMR 7.1.0부터 Amazon EMR 구성 API를 사용하여 추가 시스템 지표를 사용하고, 애플리케이션 지표를 추가하며, 지표 대상을 변경하도록 Amazon CloudWatch 에이전트를 구성할 수 있습니다. EMR 구성 API를 사용하여 클러스터의 애플리케이션을 구성하는 방법에 대한 자세한 내용은 애플리케이션 구성을 참조하세요.

> **Note**
>
> 7.1.0은 재구성 유형 `OVERWRITE`만 지원합니다. 재구성 유형에 대한 자세한 내용은 인스턴스 그룹을 재구성할 때 고려 사항을 참조하세요.

주제

- 구성 스키마
- 시스템 지표 구성 예제
- 애플리케이션 지표 구성 예제
- Amazon Managed Service for Prometheus 예제

구성 스키마

emr-metrics에는 다음과 같은 분류가 있습니다.

- emr-system-metrics - CPU, 디스크 및 메모리와 같은 시스템 지표 구성.
- emr-hadoop-hdfs-datanode-metrics - Hadoop DataNode JMX 지표 구성
- emr-hadoop-hdfs-namenode-metrics - Hadoop NameNode JMX 지표 구성
- emr-hadoop-yarn-nodemanager-metrics - Yarn NodeManager JMX 지표 구성
- emr-hadoop-yarn-resourcemanager-metrics - Yarn ResourceManager JMX 지표 구성
- emr-hbase-master-metrics - HBase Master JMX 지표 구성
- emr-hbase-region-server-metrics - HBase 리전 서버 JMX 지표 구성
- emr-hbase-rest-server-metrics - HBase REST 서버 JMX 지표 구성
- emr-hbase-thrift-server-metrics - HBase Thrift 서버 JMX 지표 구성

다음 표에서는 모든 분류에 대해 사용 가능한 속성 및 구성을 설명합니다.

emr-metrics 속성

속성	필수	설명	기본값	가능한 값	Notes
metrics_destination	선택 사항	클러스터 지표가 Amazon CloudWatch 또는 Amazon Managed Service for Prometheus 에 게시되는지 여부를 결정합니다.	"CLOUDWATCH"	"CLOUDWATCH", "PROMETHEUS"	이 속성은 대소문자를 구분하지 않습니다. 예를 들어 "Cloudwatch"는 "CLOUDWATCH"와 동일합니다.
prometheus_endpoint	선택 사항	metrics_destination 이 "PROMETHE	N/A	유효한 Amazon Managed Service for	metrics_destination 이 "PROMETHE

속성	필수	설명	기본값	가능한 값	Notes
		US"로 설정된 경우 이 속성은 제공된 Amazon Managed Service for Prometheus 원격 쓰기 엔드포인트로 지표를 전송하도록 CloudWatch 에이전트를 구성합니다.		Prometheus 원격 쓰기 URL. 원격 쓰기 URL 형식은 다음과 같습니다. `https://aps-workspaces. <region> s.com/workspaces/ <workspace_id> /api/v1/remote_write`	US"로 설정된 경우 이 필드는 필수입니다. 키를 제공하지 않거나 값이 빈 문자열인 경우 프로비저닝이 실패합니다.

emr-system-metrics 속성

속성	필수	설명	기본값	가능한 값	Notes
`metrics_collection_interval`	선택 사항	CloudWatch 에이전트에서 지표를 수집 및 게시하는 빈도(초 단위).	"60"	초 수를 지정하는 문자열. 정수만 수락합니다.	개별 지표 그룹의 속성으로 이 `metrics_collection_interval` 속성을 재정의할 수 있습니다.

emr-system-metrics 구성

cpu

속성	필수	설명	기본값	가능한 값	Notes
metrics	선택 사항	에이전트에서 수집할 CPU 지표의 목록.	Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표 참조	cpu_ 접두사(예: usage_active 및 cpu_time_idle)가 포함되거나 포함되지 않은 유효한 CPU 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	빈 문자열을 지정하면 CPU 지표를 게시하지 않습니다.
metrics_collection_interval	선택 사항	에이전트가 CPU 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 metrics_collection_interval 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 CPU 지표에 대해서만 전역 metrics_collection_interval 속성을 재정의합니다.
drop_original_metrics	선택 사항	집계되지 않은 지표를 게시하지 않을 CPU 지표 목록.	집계되지 않은 CPU 지표는 게시되지 않았습니다.	지표 속성에도 지정된 CPU 지표의 쉼표로 구분된 목록. 빈	CloudWatch 에이전트는 클러스터 ID, 인스턴스 ID, 노드 유

속성	필수	설명	기본값	가능한 값	Notes
				문자열인 경우 모든 CPU 지표를 게시합니다.	형 및 서비스 이름별로 모든 지표를 집계합니다. 기본적으로 CloudWatch 에이전트는 여러 리소스가 있는 지표에 대해 리소스당 지표를 게시하지 않습니다.
`resources`	선택 사항	에이전트가 코어당 지표를 게시할지 여부를 결정합니다.	"*"	"*"는 코어당 지표를 활성화합니다. ""는 코어당 지표를 비활성화합니다.	CloudWatch 에이전트는 `drop_original_metrics` 에서 삭제되지 않은 CPU 지표에 대한 코어당 지표만 게시합니다.

disk

속성	필수	설명	기본값	가능한 값	Notes
`metrics`	선택 사항	에이전트에서 수집할 디스크 지표의 목록.	[Amazon EMR을 사용하는 CloudWatch 에이전트의]()	`disk_` 접두사(예: `disk_total` 및 `used_perc`	빈 문자열을 지정하면 디스크 지표를 게시하지 않습니다.

속성	필수	설명	기본값	가능한 값	Notes
			기본 지표 참조	ent)가 포함되거나 포함되지 않은 유효한 디스크 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	
metrics_collection_interval	선택 사항	에이전트가 디스크 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 metrics_collection_interval 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 디스크 지표에 대해서만 전역 metrics_collection_interval 속성을 재정의합니다.

속성	필수	설명	기본값	가능한 값	Notes
`drop_original_metrics`	선택 사항	집계되지 않은 지표를 게시하지 않을 디스크 지표의 목록.	집계되지 않은 디스크 지표는 게시되지 않았습니다.	지표 속성에도 지정된 디스크 지표의 쉼표로 구분된 목록. 빈 문자열인 경우 모든 디스크 지표를 게시합니다.	CloudWatch 에이전트는 클러스터 ID, 인스턴스 ID, 노드 유형 및 서비스 이름별로 모든 지표를 집계합니다. 기본적으로 CloudWatch 에이전트는 여러 리소스가 있는 지표에 대해 리소스당 지표를 게시하지 않습니다.
`resources`	선택 사항	에이전트가 탑재 포인트당 지표를 게시할지 여부를 결정합니다.	"*"	"*"는 모든 탑재 지점을 의미하고, ""는 탑재 지점이 없거나 쉼표로 구분된 탑재 지점 목록을 의미합니다. 예를 들어 `"/,/emr"`입니다.	CloudWatch 에이전트는 `drop_original_metrics` 에서 삭제되지 않은 디스크 지표에 대한 탑재 지점당 지표만 게시합니다.

diskio

속성	필수	설명	기본값	가능한 값	Notes
metrics	선택 사항	에이전트에서 수집할 디스크 IO 지표의 목록.	Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표 참조	diskio_ 접두사(예: diskio_reads 및 writes)가 포함되거나 포함되지 않은 유효한 디스크 IO 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	빈 문자열을 지정하면 디스크 IO 지표를 게시하지 않습니다.
metrics_collection_interval	선택 사항	에이전트가 디스크 IO 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 metrics_collection_interval 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 디스크 IO 지표에 대해서만 전역 metrics_collection_interval 속성을 재정의합니다.
drop_original_metrics	선택 사항	집계되지 않은 지표를 게시하지 않을	집계되지 않은 디스크 IO 지표는 게시	지표 속성에도 지정된 디스크 IO 지표의 쉼표로 구	CloudWatch 에이전트는 클러스터 ID, 인스턴스

속성	필수	설명	기본값	가능한 값	Notes
		디스크 IO 지표의 목록.	되지 않았습니다.	분된 목록. 빈 문자열인 경우 모든 디스크 IO 지표를 게시합니다.	ID, 노드 유형 및 서비스 이름별로 모든 지표를 집계합니다. 기본적으로 CloudWatch 에이전트는 여러 리소스가 있는 지표에 대해 리소스당 지표를 게시하지 않습니다.
`resources`	선택 사항	에이전트가 디바이스당 지표를 게시할지 여부를 결정합니다.	"*"	"*"는 모든 스토리지 디바이스를 의미하며, ""는 스토리지 디바이스가 없거나 쉼표로 구분된 디바이스 이름 목록을 의미합니다. 예를 들어 `"nvme0n1, nvme1n1"` 입니다.	CloudWatch 에이전트는 `drop_original_metrics` 에서 삭제되지 않은 디스크 IO 지표에 대한 디바이스당 지표만 게시합니다.

mem

속성	필수	설명	기본값	가능한 값	Notes
`metrics`	선택 사항	에이전트에서 수집할 메모리 지표의 목록.	Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표 참조	mem_ 접두사(예: `mem_available` 및 `available_percent`)가 포함되거나 포함되지 않은 유효한 메모리 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	빈 문자열을 지정하면 메모리 지표를 게시하지 않습니다.
`metrics_collection_interval`	선택 사항	에이전트가 메모리 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 `metrics_collection_interval` 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 메모리 지표에 대해서만 전역 `metrics_collection_interval` 속성을 재정의합니다.

net

속성	필수	설명	기본값	가능한 값	Notes
metrics	선택 사항	에이전트에서 수집할 네트워크 지표의 목록.	Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표 참조	net_ 접두사(예: net_packets_sent 및 packets_recv)가 포함되거나 포함되지 않은 유효한 네트워크 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	빈 문자열을 지정하면 네트워크 지표를 게시하지 않습니다.
metrics_collection_interval	선택 사항	에이전트가 네트워크 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 metrics_collection_interval 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 네트워크 지표에 대해서만 전역 metrics_collection_interval 속성을 재정의합니다.

속성	필수	설명	기본값	가능한 값	Notes
`drop_original_metrics`	선택 사항	집계되지 않은 지표를 게시하지 않을 네트워크 지표 목록.	집계되지 않은 네트워크 지표는 게시되지 않았습니다.	지표 속성에도 지정된 네트워크 지표의 쉼표로 구분된 목록. 빈 문자열인 경우 모든 네트워크 지표를 게시합니다.	CloudWatch 에이전트는 클러스터 ID, 인스턴스 ID, 노드 유형 및 서비스 이름별로 모든 지표를 집계합니다. 기본적으로 CloudWatch 에이전트는 여러 리소스가 있는 지표에 대해 리소스당 지표를 게시하지 않습니다.
`resources`	선택 사항	에이전트가 인터페이스당 지표를 게시할지 여부를 결정합니다.	"*"	"*"는 모든 네트워크 인터페이스를 의미하며, ""는 네트워크 인터페이스가 없거나 쉼표로 구분된 인터페이스 이름 목록을 의미합니다. 예를 들어 `"eth0,eth1"` 입니다.	CloudWatch 에이전트는 `drop_original_metrics` 에서 삭제되지 않은 네트워크 지표에 대한 인터페이스당 지표만 게시합니다.

netstat

속성	필수	설명	기본값	가능한 값	Notes
metrics	선택 사항	에이전트에서 수집할 네트워크 통계 지표의 목록.	Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표 참조	netstat_ 접두사(예: tcp_listen 및 netstat_udp_socket)가 포함되거나 포함되지 않은 유효한 메모리 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	빈 문자열을 지정하면 네트워크 통계 지표를 게시하지 않습니다.
metrics_collection_interval	선택 사항	에이전트가 네트워크 통계 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 metrics_collection_interval 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 네트워크 통계 지표에 대해서만 전역 metrics_collection_interval 속성을 재정의합니다.

processes

속성	필수	설명	기본값	가능한 값	Notes
metrics	선택 사항	에이전트에서 수집할 프로세스 지표의 목록.	Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표 참조	processes _ 접두사(예: processes_running 및 total)가 포함되거나 포함되지 않은 유효한 메모리 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	빈 문자열을 지정하면 프로세스 지표를 게시하지 않습니다.
metrics_collection_interval	선택 사항	에이전트가 시스템 프로세스 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 metrics_collection_interval 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 시스템 프로세스 지표에 대해서만 전역 metrics_collection_interval 속성을 재정의합니다.

swap

속성	필수	설명	기본값	가능한 값	Notes
`metrics`	선택 사항	에이전트에서 수집할 스왑 지표의 목록.	Amazon EMR을 사용하는 CloudWatch 에이전트의 기본 지표 참조	swap_ 접두사(예: `swap_free` 및 `used_percent`)가 포함되거나 포함되지 않은 유효한 메모리 지표 이름의 쉼표로 구분된 목록. 유효한 지표는 CloudWatch 에이전트가 수집하는 지표를 참조하세요.	빈 문자열을 지정하면 스왑 지표를 게시하지 않습니다.
`metrics_collection_interval`	선택 사항	에이전트가 스왑 지표를 수집하고 게시해야 하는 빈도(초 단위).	전역 `metrics_collection_interval` 의 값.	초 수를 지정하는 문자열. 정수만 수락합니다.	이 값은 스왑 지표에 대해서만 전역 `metrics_collection_interval` 속성을 재정의합니다.

emr-hadoop-hdfs-datanode-metrics 속성

속성	필수	설명	기본값	가능한 값
`<custom_bean_name>`	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수집해야 하는 MBean(예: `Hadoop:service=DataNode,name=DataNodeActivity`). Amazon EMR 릴리스 7.0에 대한 [JMX YAML 파일 예제]에서 샘플 MBean 이름 및 해당 지표를 찾을 수 있습니다.	MBean과 연결된 쉼표로 구분된 지표 목록을 포함하는 문자열. 예를 들어 `BlocksCached,BlocksRead` 입니다.
`otel.metric.export.interval`	선택 사항	Hadoop DataNode 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

emr-hadoop-hdfs-namenode-metrics 속성

속성	필수	설명	기본값	가능한 값
`<custom_bean_name>`	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수	MBean과 연결된 쉼표로 구분된 지

속성	필수	설명	기본값	가능한 값
			집해야 하는 MBean(예: `Hadoop:service=NameNode,name=FSNamesystem`). Amazon EMR 릴리스 7.0에 대한 [JMX YAML 파일 예제](#)에서 샘플 MBean 이름 및 해당 지표를 찾을 수 있습니다.	표 목록을 포함하는 문자열. 예를 들어 `BlockCapacity,CapacityUsedGB` 입니다.
`otel.metric.export.interval`	선택 사항	Hadoop NameNode 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

emr-hadoop-yarn-nodemanager-metrics 속성

속성	필수	설명	기본값	가능한 값
<custom_bean_name>	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수집해야 하는 MBean(예: `Hadoop:service=Nod`	MBean과 연결된 쉼표로 구분된 지표 목록을 포함하는 문자열. 예를 들어 `MaxCapaci`

속성	필수	설명	기본값	가능한 값
			eManager, name=NodeManagerMetrics). Amazon EMR 릴리스 7.0 에 대한 JMX YAML 파일 예제에서 샘플 MBean 이름 및 해당 지표를 찾을 수 있습니다.	ty,AllocatedGB 입니다.
`otel.metric.export.interval`	선택 사항	Hadoop YARN NodeManager 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

emr-hadoop-yarn-resourcemanager-metrics 속성

속성	필수	설명	기본값	가능한 값
<custom_bean_name>	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수집해야 하는 MBean(예: Hadoop:service=ResourceMana	MBean과 연결된 쉼표로 구분된 지표 목록을 포함하는 문자열. 예를 들어 MaxCapacity,MaxCap

속성	필수	설명	기본값	가능한 값
			ger,name= PartitionQueueMetrics). Amazon EMR 릴리스 7.0 에 대한 JMX YAML 파일 예제에서 샘플 MBean 이름 및 해당 지표를 찾을 수 있습니다.	acityVCores 입니다.
otel.metric.export.interval	선택 사항	Hadoop YARN ResourceManager 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

emr-hbase-master-metrics 속성

속성	필수	설명	기본값	가능한 값
<custom_bean_name>	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수집해야 하는 MBean(예: Hadoop:service=HBase,name=M	MBean과 연결된 쉼표로 구분된 지표 목록을 포함하는 문자열. 예를 들어 AssignFailedCount,

속성	필수	설명	기본값	가능한 값
			aster,sub =AssignmentManager). Amazon EMR 릴리스 7.0에 대한 [JMX YAML 파일 예제](#)에서 샘플 MBean 이름 및 해당 지표를 찾을 수 있습니다.	AssignSubmittedCount 입니다.
otel.metric.export.interval	선택 사항	HBase Master 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

emr-hbase-region-server-metrics 속성

속성	필수	설명	기본값	가능한 값
<custom_bean_name>	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수집해야 하는 MBean(예: Hadoop:service=HBase,name=RegionServer,sub=IP	MBean과 연결된 쉼표로 구분된 지표 목록을 포함하는 문자열. 예를 들어 numActiveHandler,numActivePriorityHa

속성	필수	설명	기본값	가능한 값
			C). Amazon EMR 릴리스 7.0에 대한 JMX YAML 파일 예제에서 샘플 MBean 이름 및 해당 지표를 찾을 수 있습니다.	ndler 입니다.
otel.metric.export.interval	선택 사항	HBase 리전 서버 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

emr-hbase-rest-server-metrics 속성

속성	필수	설명	기본값	가능한 값
<custom_bean_name>	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수집해야 하는 MBean(예: `Hadoop:service=HBase,name=REST`). Amazon EMR 릴리스 7.0에 대한 JMX YAML 파일 예제에서 샘	MBean과 연결된 쉼표로 구분된 지표 목록을 포함하는 문자열. 예를 들어 `successfulPut,successfulScanCount` 입니다.

속성	필수	설명	기본값	가능한 값
			플 MBean 이름 및 해당 지표를 찾을 수 있습니다.	
otel.metric.export.interval	선택 사항	HBase Rest Server 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

emr-hbase-thrift-server-metrics 속성

속성	필수	설명	기본값	가능한 값
<custom_bean_name>	선택 사항	N/A	CloudWatch 에이전트에서 지표를 수집해야 하는 MBean(예: `Hadoop:service=HBase,name=Thrift,sub=ThriftOne`). Amazon EMR 릴리스 7.0에 대한 [JMX YAML 파일 예제]에서 샘플 MBean 이름 및 해당 지	MBean과 연결된 쉼표로 구분된 지표 목록을 포함하는 문자열. 예를 들어 `BatchGet_max,BatchGet_mean` 입니다.

속성	필수	설명	기본값	가능한 값
			표를 찾을 수 있습니다.	
otel.metric.export.interval	선택 사항	HBase Thrift 서버 지표를 수집하는 빈도(밀리초 단위).	"60000"	밀리초 수를 지정하는 문자열. 정수만 수락합니다.

시스템 지표 구성 예제

다음 예제에서는 모든 시스템 지표 내보내기를 중지하도록 CloudWatch 에이전트를 구성하는 방법을 보여줍니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

다음 예제에서는 기본 시스템 지표를 내보내도록 CloudWatch 에이전트를 구성합니다. 그러면 시스템 지표를 이미 한 번 이상 재구성한 경우 기본 시스템 지표 내보내기만 가능하도록 에이전트를 다시 빠르게 재설정할 수 있습니다. 이 재설정은 이전에 재구성된 모든 애플리케이션 지표도 제거합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": []
  }
```

```
]
```

다음 예제에서는 cpu, mem 및 disk 지표를 내보내도록 클러스터를 구성합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-system-metrics",
        "Properties": {
          "metrics_collection_interval": "20"
        },
        "Configurations": [
          {
            "Classification": "cpu",
            "Properties": {
              "metrics": "cpu_usage_guest,cpu_usage_idle",
              "metrics_collection_interval": "30",
              "drop_original_metrics": "cpu_usage_guest"
            }
          },
          {
            "Classification": "mem",
            "Properties": {
              "metrics": "mem_active"
            }
          },
           {
            "Classification": "disk",
            "Properties": {
              "metrics": "disk_used_percent",
              "resources": "/,/mnt",
              "drop_original_metrics": ""
            }
          }
        ]
      }
    ]
  }
]
```

이전 예제 구성에는 다음 속성이 있습니다.

- 에이전트는 30초마다 모든 CPU에 대해 `cpu_guest` 지표를 수집합니다. CloudWatch 네임스페이스 `CWAgent > cluster.id, instance.id, node.type, service.name`에서 집계된 지표를 찾을 수 있습니다.
- 에이전트는 30초마다 모든 CPU에 대해 `cpu_idle` 지표를 수집합니다. CloudWatch 네임스페이스 `CWAgent > cluster.id, instance.id, node.type, service.name`에서 집계된 지표를 찾을 수 있습니다. 에이전트는 CPU당 지표도 수집합니다. 동일한 네임스페이스에서 이를 찾을 수 있습니다. `drop_original_metrics` 속성에 `cpu_idle`가 포함되어 있지 않으므로 에이전트는 이 지표를 수집합니다. 에이전트는 지표를 무시하지 않습니다.
- 에이전트는 20초마다 `mem_active` 지표를 수집합니다. CloudWatch 네임스페이스 `CWAgent > cluster.id, instance.id, node.type, service.name`에서 집계된 지표를 찾을 수 있습니다.
- 에이전트는 20초마다 `/` 및 `/mnt` 디스크 탑재에 대한 `disk_used_percent` 지표를 수집합니다. CloudWatch 네임스페이스 `CWAgent > cluster.id, instance.id, node.type, service.name`에서 집계된 지표를 찾을 수 있습니다. 에이전트는 탑재당 지표도 수집합니다. 동일한 네임스페이스에서 이를 찾을 수 있습니다. `drop_original_metrics` 속성에 `disk_used_percent`가 포함되어 있지 않으므로 에이전트는 이 지표를 수집합니다. 에이전트는 지표를 무시하지 않습니다.

애플리케이션 지표 구성 예제

다음 예제에서는 Hadoop Namenode 서비스에 대한 지표 내보내기를 중지하도록 CloudWatch 에이전트를 구성합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {},
        "Configurations": []
      }
    ]
  }
]
```

다음 예제에서는 Hadoop 애플리케이션 지표를 내보내도록 클러스터를 구성합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-namenode-metrics",
        "Properties": {
          "Hadoop:service=NameNode,name=FSNamesystem": "BlockCapacity,CapacityUsedGB",
          "otel.metric.export.interval": "20000"
        },
        "Configurations": []
      },
       {
        "Classification": "emr-hadoop-hdfs-datanode-metrics",
        "Properties": {
          "Hadoop:service=DataNode,name=JvmMetrics": "MemNonHeapUsedM",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      },
       {
        "Classification": "emr-hadoop-yarn-resourcemanager-metrics",
        "Properties": {
          "Hadoop:service=ResourceManager,name=CapacitySchedulerMetrics":
 "AllocateNumOps,NodeUpdateNumOps"
        },
        "Configurations": []
      }
    ]
  }
]
```

이전 예제에는 다음 속성이 있습니다.

- 에이전트는 20초마다 Hadoop Namenode 서비스를 실행하는 인스턴스에서 `BlockCapacity` 및 `CapacityUsedGB` 지표를 수집합니다.
- 에이전트는 30초마다 Hadoop Datanode 서비스를 실행하는 인스턴스에서 `MemNonHeapUsedM` 지표를 수집합니다.

• 에이전트는 30초마다 Hadoop YARN ResourceManager 실행하는 인스턴스에서 `AllocateNumOps` 및 `NodeUpdateNumOps` 지표를 수집합니다.

Amazon Managed Service for Prometheus 예제

다음 예제에서는 Amazon Managed Service for Prometheus로 지표를 전송하도록 CloudWatch 에이전트를 구성하는 방법을 보여줍니다.

현재 Amazon Managed Service for Prometheus로 지표를 내보내고 있으며, 클러스터의 지표를 재구성하여 Amazon Managed Service for Prometheus로 지표를 계속 내보내려는 경우 `metrics_destination` 및 `prometheus_endpoint` 속성을 포함해야 합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "http://amp-workspace/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

CloudWatch 에이전트를 사용하여 지표를 CloudWatch로 내보내려면 다음 예제를 사용합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

> **Note**
>
> CloudWatch 에이전트에는 특정 속성의 이름을 바꾸는 Prometheus 익스포터가 있습니다. Amazon Managed Service for Prometheus는 기본 지표 레이블의 경우 Amazon CloudWatch에서 사용하는 기간 대신 밑줄 문자를 사용합니다. Amazon Managed Grafana를 사용하

여 Amazon Managed Service for Prometheus에서 기본 지표를 시각화하는 경우 레이블은 `cluster_id`, `instance_id`, `node_type` 및 `service_name`으로 표시됩니다.

Amazon EMR 7.0.0용 CloudWatch 에이전트 구성

기본 CloudWatch 에이전트 구성이 제공하는 지표 외에 추가 시스템 지표를 사용하도록 Amazon CloudWatch 에이전트를 구성할 수 있습니다. 7.0.0을 구성하려면 부트스트랩 작업을 사용해야 합니다. 부트스트랩 작업에 대한 예시는 다음 섹션에 나와 있습니다. 향후 릴리스에서 Amazon EMR은 Amazon EMR API를 통해 추가 구성 옵션을 제공할 예정입니다.

주제

Amazon EMR 7.0.0을 사용하여 추가 시스템 지표를 구성하세요.

Amazon EMR 7.0.0에서 다른 시스템 지표 세트를 사용하도록 에이전트를 구성하려면 다음 단계를 사용하세요.

1. Amazon S3 계정에서 CloudWatch 에이전트 지표를 지정하는 구성 파일을 저장할 버킷을 만들거나 선택합니다.
2. 선호하는 지표를 지정하여 `emr-amazon-cloudwatch-agent.json` 구성 파일을 생성합니다. 이렇게 하려면 CloudWatch 에이전트 구성 파일 생성에서 설명된 방법 중 하나를 사용하세요. CloudWatch 에이전트 구성 파일의 구조에 대한 자세한 내용은 Amazon CloudWatch 사용 설명서의 CloudWatch 에이전트 구성 파일 수동 생성 또는 편집을 참조하세요.
3. 다음으로 GitHub의 aws-emr-utilities 리포지토리로 이동하여 다음 시스템 지표 스크립트를 다운로드합니다.
   - **`install_system_metrics_launcher.sh`** – 노드가 부트스트래핑을 완료할 수 있도록 다운로드한 다음 백그라운드에서 `install_system_metrics.sh`를 실행하는 스크립트입니다.
   - **`install_system_metrics.sh`** – 실행되는 인스턴스가 부트스트래핑을 마칠 때까지 기다린 다음 구성을 JSON 파일로 다운로드하여 적용하는 스크립트입니다.
4. 각 SH 파일을 열고 *`amzn-s3-demo-bucket`*을 1단계의 버킷 이름으로 바꿉니다.

5. JSON과 SH 파일 두 개를 S3 버킷에 업로드합니다.
6. 이제 Amazon EMR 콘솔로 이동하고 CloudWatch 에이전트를 사용하여 새 클러스터를 생성할 수 있습니다. 왼쪽 탐색 창의 EMR on EC2에서 클러스터를 선택한 다음 클러스터 생성을 선택합니다.
7. 이름 및 애플리케이션 섹션에서 Amazon EMR 릴리스 7.0.0 이상을 선택합니다.
8. 애플리케이션 번들에서 클러스터에 설치할 번들 또는 앱을 선택하고 선택 항목에 Amazon CloudWatch 에이전트를 포함합니다.
9. 부트스트랩 작업 섹션에서 추가를 선택합니다.
   - 이름에 **install_system_metrics_launcher.sh**를 삽입합니다.
   - 스크립트 위치에 **s3://*amzn-s3-demo-bucket*/install_system_metrics_launcher.sh**을 삽입하세요. *amzn-s3-demo-bucket*을 S3 버킷의 경로로 바꾸세요.
   - 인수 블록은 비워 둡니다.
10. 부트스트랩 작업 추가를 선택합니다.
11. 워크로드 요구 사항에 맞게 클러스터를 계속 생성하세요.

클러스터가 시작되면 CloudWatch 에이전트는 구성 파일에 지정한 시스템 지표를 CloudWatch에 게시합니다.

Amazon EMR 7.0.0을 사용하여 애플리케이션 지표를 구성합니다.

시스템 지표 외에도 HDFS 및 YARN에 대한 애플리케이션 지표를 게시하도록 Amazon CloudWatch 에이전트를 구성할 수 있습니다. 다음 단계에 따라 에이전트가 애플리케이션 지표를 게시하도록 구성합니다.

1. Amazon S3 계정에서 CloudWatch 에이전트 지표를 지정하는 구성 파일을 저장할 버킷을 만들거나 선택합니다.
2. 다음으로 GitHub의 aws-emr-utilities 리포지토리로 이동하여 다음 스크립트를 다운로드합니다.
   - install_app_metrics_launcher.sh – 노드가 부트스트래핑을 완료할 수 있도록 다운로드한 다음 백그라운드에서 install_app_metrics.sh를 실행하는 스크립트입니다.
   - install_app_metrics.sh – 실행되는 인스턴스가 부트스트래핑을 완료할 때까지 기다린 후 다음 단계에서 다운로드할 YAML 파일의 구성을 다운로드하고 적용하는 스크립트입니다.
3. 각 파일을 열고 *amzn-s3-demo-bucket*을 1단계의 버킷 이름으로 바꿉니다.

4. 다음으로 다음 YAML 매핑 파일을 다운로드합니다. 이러한 YAML 파일의 구조에 대한 자세한 내용은 OpenTelemetry Instrumentation for JavaGitHub 리포지토리의 `javaagent` 섹션을 참조하세요.

   - **`datanode-metrics.yaml`** – Hadoop DataNode 지표의 구성입니다.
   - **`namenode-metrics.yaml`** – Hadoop NameNode 지표의 구성입니다.
   - **`nodemanager-metrics.yaml`** – Yarn NodeManager 지표의 구성입니다.
   - **`resourcemanager-metrics.yaml`** – Yarn ResourceManager 지표의 구성입니다.
5. S3 버킷에 SH 파일 2개와 YAML 파일 4개를 업로드합니다.
6. 이제 Amazon EMR 콘솔로 이동하고 CloudWatch 에이전트를 사용하여 새 클러스터를 생성할 수 있습니다. 왼쪽 탐색 창의 EMR on EC2에서 클러스터를 선택한 다음 클러스터 생성을 선택합니다.
7. 이름 및 애플리케이션 섹션에서 Amazon EMR 릴리스 7.0.0 이상을 선택합니다.
8. 애플리케이션 번들에서 클러스터에 설치할 번들 또는 앱의 사용자 지정 그룹을 선택하고 선택 항목에 CloudWatch 에이전트를 포함합니다.
9. 부트스트랩 작업 섹션에서 추가를 선택합니다.

   - 이름에 **`install_app_metrics_launcher.sh`**를 삽입합니다.
   - 스크립트 위치에 **`s3://`*`amzn-s3-demo-bucket`*`/install_app_metrics_launcher.sh`**을 삽입하세요. *`amzn-s3-demo-bucket`*을 S3 버킷의 경로로 바꾸세요.
   - 인수 블록은 비워 둡니다.
10. 부트스트랩 작업 추가를 선택합니다.
11. 워크로드 요구 사항에 맞게 클러스터를 계속 생성하세요.

클러스터가 시작되면 CloudWatch 에이전트는 시스템 지표와 함께 지정한 애플리케이션 지표를 CloudWatch에 게시합니다.

Amazon EMR 7.0.0을 사용하여 Amazon Managed Service for Prometheus를 지표용 클라우드 스토리지로 구성합니다.

Amazon CloudWatch 에이전트에서 CloudWatch 대신 Amazon Managed Service for Prometheus에 지표를 게시하도록 구성할 수 있습니다.

> **Note**
>
> Amazon CloudWatch 에이전트의 지표를 Amazon Managed Service for Prometheus 또는 Amazon CloudWatch에 게시할 수 있지만, 동일한 클러스터의 두 서비스에 지표를 게시할 수는 없습니다.

Amazon Managed Service for Prometheus에 지표를 게시하도록 에이전트를 구성하려면 Amazon EMR의 Amazon EC2 인스턴스 프로파일에 `aps:RemoteWrite` AWS Identity and Access Management (IAM) 권한을 추가해야 합니다. 다음 예제 정책은 필요한 권한을 포함합니다.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": "aps:RemoteWrite",
      "Resource": "*"
    }
  ]
}
```

EMR 클러스터의 CloudWatch 에이전트를 사용하여 Amazon Managed Service for Prometheus에 지표를 게시합니다.

서비스 정책에 올바른 권한이 부여되면 다음 단계를 사용하여 CloudWatch 에이전트를 사용하여 Amazon Managed Service for Prometheus에 지표를 게시하는 클러스터를 시작합니다.

1. AWS Management Console 또는를 사용하여 Amazon Managed Service for Prometheus 워크스페이스를 AWS CLI 생성합니다. 자세한 내용은 Amazon Managed Service for Prometheus 사용 설명서의 작업 영역 생성을 참조하세요.
2. Amazon S3 계정에서 Amazon Managed Service for Prometheus를 클라우드 스토리지로 지정하는 시작 파일을 저장할 버킷을 만들거나 선택합니다.
3. 다음으로 GitHub의 aws-emr-utilities 리포지토리로 이동하여 다음 스크립트를 다운로드합니다.

- **add_prometheus_endpoint_launcher.sh** – 노드가 부트스트래핑을 완료할 수 있도록 다운로드한 다음 백그라운드에서 add_prometheus_endpoint.sh를 실행하는 스크립트입니다.
- **add_prometheus_endpoint.sh** – 실행되는 인스턴스가 부트스트래핑을 마칠 때까지 기다린 다음 클러스터를 시작할 때 인수로 제공하는 Amazon Managed Service for Prometheus 엔드포인트에 게시하도록 CloudWatch 에이전트를 구성하는 스크립트입니다.

4. 각 파일을 열고 *amzn-s3-demo-bucket*을 2단계의 버킷 이름으로 바꿉니다.
5. AWS CLI 를 사용하여 add_prometheus_endpoint_launcher.sh 부트스트랩 작업으로 EMR 클러스터를 생성합니다. 다음 명령에서 *amzn-s3-demo-bucket*을 부트스트랩 작업을 보관하는 버킷으로 교체하고 *managedpro-remote-write-workspace-url*를 Amazon Managed Service for Prometheus 작업 공간의 원격 쓰기 엔드포인트로 대체합니다. Amazon EMR 출시 레이블을 다음 emr-7.0.0 또는 그 이상으로 지정해야 합니다.

```
aws emr create-cluster --name managedpro-cluster \
    --release-label emr-7.0.0 \
    --applications Name=Hadoop Name=AmazonCloudWatchAgent \
    --ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
    --instance-count 3 --use-default-roles
   --bootstrap-actions Name='Add Prometheus Endpoint',Path=s3://amzn-s3-demo-bucket/add_prometheus_endpoint_launcher.sh,Args='managedpro-remote-write-workspace-url'
```

클러스터가 시작되면 CloudWatch 에이전트는 수집한 지표를 Amazon Managed Service for Prometheus에 게시합니다.

Amazon Managed Service for Prometheus를 Amazon Managed Grafana의 데이터 소스로 사용

Amazon EMR이 Amazon Managed Service for Prometheus에 클러스터 지표를 게시한 후에는 다음 단계를 사용하여 Amazon Managed Grafana로 지표를 시각화할 수 있습니다.

1. AWS Management Console 를 사용하여 적절한 권한이 있는 Amazon Managed Grafana 워크스페이스 및 사용자를 생성합니다. 자세한 내용은 Amazon Managed Grafana 사용 설명서의 작업 공간 생성을 참조하세요.
2. Amazon Managed Service for Prometheus 작업 영역을 Amazon Managed Grafana에 데이터 소스로 추가합니다. 자세한 내용은 Amazon Managed Grafana 사용 설명서에서 AWS 데이터 소스 구성을 사용하여 Amazon Managed Service for Prometheus를 데이터 소스로 추가를 참조하세요.

> **Note**
>
> CloudWatch 에이전트에는 특정 속성의 이름을 바꾸는 Prometheus 익스포터가 있습니다. Amazon Managed Service for Prometheus는 기본 지표 레이블의 경우 Amazon CloudWatch에서 사용하는 기간 대신 밑줄 문자를 사용합니다. 따라서 Amazon Managed Grafana를 사용하여 Amazon Managed Service for Prometheus에서 기본 지표를 시각화하는 경우 레이블은 `jobflow_id`, `instance_id` 및 `service_name`으로 표시됩니다.
> 또한 CloudWatch 에이전트가 Amazon Managed Service for Prometheus에 게시하는 모든 애플리케이션 지표에는 `service_name` 대신 `job` 레이블이 사용됩니다. 하지만 시스템 지표에는 계속해서 이 `service_name` 레이블이 사용됩니다.

고려 사항 및 제한 사항

- Amazon EMR 릴리스 7.0.0 이상을 사용하여 생성한 클러스터에 네이티브 Amazon CloudWatch 에이전트를 사용할 수 있습니다.
- Amazon EMR 7.0.0에서 CloudWatch 에이전트를 구성하려면 [부트스트랩 작업]을 사용해야 합니다. 향후 릴리스에서 Amazon EMR은 Amazon EMR API를 통해 추가 구성 옵션을 제공할 예정입니다.
- 클러스터를 생성하는 리전에서 에이전트(SSM 에이전트)와 같은 다른 방법으로 CloudWatch AWS Systems Manager 에이전트를 이미 배포한 경우 Amazon EMR CloudWatch 에이전트를 설치할 수 없습니다. CloudWatch 이는 오류를 유발하며 클러스터가 종료됩니다.
- CloudWatch GetMetricData API는 요청당 최대 500개의 지표를 지원합니다. Amazon EMR 클러스터에 인스턴스 그룹 또는 플릿에 250개 이상의 노드가 포함된 경우 EMR의 CloudWatch 임베디드 대시보드에 있는 해당 그래프는 클러스터 개요 대시보드에서 지표당 2개의 데이터 포인트가 필요하므로 지표가 너무 많음 오류와 함께 공백으로 표시됩니다. 그러나 코어 또는 태스크 인스턴스 그룹(또는 플릿) 대시보드를 필터링하면 인스턴스 그룹 또는 플릿당 최대 500개의 노드에 대한 그래프를 볼 수 있습니다. 이는 지표당 두 개의 데이터 포인트가 필요하지 않기 때문입니다. 인스턴스 그룹 또는 플릿당 500개 이상의 노드에서 이러한 대시보드의 지표에 대해 지표 과다 오류도 발생합니다.

CloudWatch 에이전트 출시 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 AmazonCloudWatchAgent의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 [Amazon EMR 7.x 릴리스 버전], [Amazon EMR 6.x 릴리스 버전] 또는 [Amazon EMR 5.x 릴리스 버전]의 릴리스의 구성 요소 버전 섹션을 참조하세요.

AmazonCloudWatchAgent 버전 정보

Amazon EMR 릴리스 레이블	AmazonCloudWatchAgent 버전	AmazonCloudWatchAgent를 사용하여 설치한 구성 요소
emr-7.9.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7.8.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd fs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server
emr-7.7.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hd

Amazon EMR 릴리스 레이블	AmazonCloudWatchAgent 버전	AmazonCloudWatchAgent를 사용하여 설치한 구성 요소
		fs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.6.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	AmazonCloudWatchAgent 버전	AmazonCloudWatchAgent를 사용하여 설치한 구성 요소
emr-7.5.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.4.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	AmazonCloudWatchAgent 버전	AmazonCloudWatchAgent를 사용하여 설치한 구성 요소
emr-7.3.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.2.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent , emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	AmazonCloudWatchAgent 버전	AmazonCloudWatchAgent를 사용하여 설치한 구성 요소
emr-7.1.0	1.300032.2	adot-java-agent, emrfs, emr-amazon-cloudwatch-agent, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.0.0	1.300031.1	adot-java-agent, emr-amazon-cloudwatch-agent

Delta Lake

Delta Lake는 일반적으로 Amazon S3에 구축되는 레이크 하우스 아키텍처를 위한 스토리지 계층 프레임워크입니다. Amazon EMR 릴리스 6.9.0 이상에서 Delta Lake 테이블을 포함하는 Amazon EMR 클러스터에서 Apache Spark 3.x를 사용할 수 있습니다. Delta Lake가 있는 레이크 하우스에 대한 자세한 내용은 https://delta.io/를 참조하세요.

다음 테이블에는 Amazon EMR이 Delta를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Delta의 버전이 나열되어 있습니다.

이 릴리스에서 Delta와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 델타 버전 정보

Amazon EMR 릴리스 레이블	Delta 버전	Delta와 함께 설치된 구성 요소
emr-7.9.0	Delta 3.3.0-amzn-1	Not available.

다음 테이블에는 Amazon EMR이 Delta를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Delta의 버전이 나열되어 있습니다.

이 릴리스에서 Delta와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Delta 버전 정보

Amazon EMR 릴리스 레이블	Delta 버전	Delta와 함께 설치된 구성 요소
emr-6.15.0	Delta 2.4.0	Not available.

Delta Lake 소개

Delta Lake는 Amazon S3에 일반적으로 구축되는 최신 데이터 레이크 아키텍처를 구현하는 데 도움이 되는 오픈 소스 프로젝트입니다. Delta Lake는 다음과 같은 기능을 제공합니다.

- Spark에서의 원자성, 일관성, 격리성, 지속성(ACID) 지원 트랜잭션. 리더는 Spark 작업 중에 테이블의 일관된 보기를 봅니다.
- Spark의 분산 처리를 통한 확장 가능한 메타데이터 처리.
- 스트리밍 및 배치 사용 사례를 동일한 Delta 테이블과 결합합니다.

- 데이터 모으기 중에 잘못된 레코드를 피하기 위한 자동 스키마 적용.
- 데이터 버전 관리를 통한 시간 이동.
- 변경 데이터 캡처(CDC), 스트리밍 업서트 등과 같은 복잡한 사용 사례에 대한 병합, 업데이트 및 삭제 작업을 지원합니다.

Delta Lake가 설치된 클러스터 사용

주제

Flink와 함께 Delta Lake 클러스터 사용

Amazon EMR 릴리스 6.11 이상에서는 Flink 클러스터에서 Delta Lake를 사용할 수 있습니다. 다음 예제에서는 AWS CLI 를 사용하여 Amazon EMR Flink 클러스터에서 Delta Lake로 작업합니다.

> **Note**
>
> Amazon EMR은 Flink 클러스터에서 Delta Lake를 사용할 때 Flink DataStream API를 지원합니다.

Delta Lake 클러스터 생성

1. 다음 콘텐츠가 포함된 `delta_configurations.json` 파일을 생성합니다.

```
[{"Classification":"delta-defaults",
    "Properties":{"delta.enabled":"true"}}]
```

2. 다음 구성을 사용하여 클러스터를 생성합니다. `example Amazon S3 bucket path` 및 `subnet ID`를 사용자 정보로 바꿉니다.

```
aws emr create-cluster
--release-label emr-6.11.0
```

```
--applications Name=Flink
--configurations file://delta_configurations.json
--region us-east-1  --name My_Spark_Delta_Cluster
--log-uri  s3://amzn-s3-demo-bucket/
--instance-type m5.xlarge
--instance-count 3
--service-role EMR_DefaultRole_V2
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

Flink yarn 세션 초기화

Flink yarn 세션을 초기화하려면 다음 명령을 실행합니다.

```
flink-yarn-session -d
```

Delta Lake를 사용하여 Flink 작업 빌드

다음 예제에서는 Delta Lake에서 sbt 또는 Maven을 사용해 Flink 작업을 빌드하는 방법을 보여줍니다.

sbt

[sbt](https://www.scala-sbt.org/)는 소규모 프로젝트의 경우 거의 또는 전혀 구성하지 않고도 사용할 수 있는 Scala용 빌드 도구입니다.

```
libraryDependencies ++= Seq(
  "io.delta" %% "delta-flink" % deltaConnectorsVersion % "provided",
  "io.delta" %% "delta-standalone" % deltaConnectorsVersion % "provided",
  "org.apache.flink" %% "flink-clients" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-parquet" % flinkVersion % "provided",
  "org.apache.hadoop" % "hadoop-client" % hadoopVersion % "provided",
  "org.apache.flink" % "flink-table-common" % flinkVersion % "provided",
  "org.apache.flink" %% "flink-table-runtime" % flinkVersion % "provided")
```

Maven

Maven은 Apache Software Foundation의 오픈 소스 빌드 자동화 도구입니다. Maven을 사용하면 Amazon EMR에서 Delta Lake를 사용하여 Flink 작업을 빌드, 게시 및 배포할 수 있습니다.

```
<project>
<properties>
```

```
    <scala.main.version>2.12</scala.main.version>
    <delta-connectors-version>0.6.0</delta-connectors-version>
    <flink-version>1.16.1</flink-version>
    <hadoop-version>3.1.0</hadoop-version>
</properties>

<dependencies>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-flink</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>io.delta</groupId>
        <artifactId>delta-standalone_$scala-main-version</artifactId>
        <version>$delta-connectors-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-clients</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-parquet</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.hadoop</groupId>
        <artifactId>hadoop-client</artifactId>
        <version>$hadoop-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
        <groupId>org.apache.flink</groupId>
        <artifactId>flink-table-common</artifactId>
        <version>$flink-version</version>
        <scope>provided</scope>
    </dependency>
    <dependency>
```

```
            <groupId>org.apache.flink</groupId>
            <artifactId>flink-table-runtime</artifactId>
            <version>$flink-version</version>
            <scope>provided</scope>
        </dependency>
    </dependencies>
```

Flink Datastream API를 사용하여 Delta 테이블에 쓰기

다음 예제를 사용하여 `deltaTablePath:`를 통해 테이블에 쓰도록 DeltaSink 생성

```
public static DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath,
        RowType rowType) {
    Configuration configuration = new Configuration();
    DeltaSink<RowData> deltaSink = DeltaSink
            .forRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration,
                    rowType)
            .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Flink Datastream API를 사용하여 Delta 테이블에서 읽기

다음 예제를 사용하여 `deltaTablePath:`를 통해 테이블에서 데이터를 읽도록 바인딩된 DeltaSource 생성

```
public static DataStream<RowData> createBoundedDeltaSourceAllColumns(
        StreamExecutionEnvironment env,
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    DeltaSource<RowData> deltaSource = DeltaSource
            .forBoundedRowData(
                    new org.apache.flink.core.fs.Path(deltaTablePath),
                    configuration)
            .build();
```

```
    return env.fromSource(deltaSource, WatermarkStrategy.noWatermarks(), "delta-
source");
}
```

Delta Lake 독립 실행형에 대한 다중 클러스터 지원을 통해 싱크 생성

다음 예제를 사용하여 deltaTablePath 및 다중 클러스터 지원을 통해 테이블에 쓰도록 DeltaSink를 생성합니다.

```
public DataStream<RowData> createDeltaSink(
        DataStream<RowData> stream,
        String deltaTablePath) {
    Configuration configuration = new Configuration();
    configuration.set("spark.delta.logStore.s3.impl",
 "io.delta.storage.S3DynamoDBLogStore");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.tableName",
 "delta_log");
    configuration.set("spark.io.delta.storage.S3DynamoDBLogStore.ddb.region", "us-
east-1");

    DeltaSink<RowData> deltaSink = DeltaSink
        .forRowData(
            new Path(deltaTablePath),
            configuration,
            rowType)
        .build();
    stream.sinkTo(deltaSink);
    return stream;
}
```

Flink 작업 실행

다음 명령을 사용하여 작업을 실행합니다.

```
flink run FlinkJob.jar
```

Trino와 함께 Delta Lake 클러스터 사용

Amazon EMR 릴리스 6.9.0 이상에서는 Trino 클러스터와 함께 Delta Lake를 사용할 수 있습니다.

이 자습서에서는를 사용하여 Amazon EMR Trino 클러스터에서 Delta Lake로 작업 AWS CLI 합니다.

Delta Lake 클러스터 생성

1. `delta_configurations.json` 파일을 생성하고 선택한 카탈로그의 값을 설정합니다. 예를 들어 Hive 메타스토어를 카탈로그로 사용하려면 파일에 다음과 같은 콘텐츠가 있어야 합니다.

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore.uri":"thrift://localhost:9083"}}]
   ```

   AWS Glue 카탈로그를 스토어로 사용하려면 파일에 다음 콘텐츠가 있어야 합니다.

   ```
   [{"Classification":"delta-defaults",
       "Properties":{"delta.enabled":"true"}},
       {"Classification":"trino-connector-delta",
       "Properties":{"hive.metastore":"glue"}}]
   ```

2. 다음과 같은 구성으로 클러스터를 생성하고 **`example Amazon S3 bucket path`** 및 **`subnet ID`**를 사용자 정보로 바꿉니다.

   ```
   aws emr create-cluster
       --release-label emr-6.9.0
       --applications Name=Trino
       --configurations file://delta_configurations.json
       --region us-east-1  --name My_Spark_Delta_Cluster
       --log-uri  s3://amzn-s3-demo-bucket/
       --instance-type m5.xlarge
       --instance-count 2
       --service-role EMR_DefaultRole_V2
       --ec2-attributes
    InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
   ```

Delta Lake용 Trino 세션 초기화

Trino 세션을 초기화하려면 다음 명령을 실행합니다.

```
trino-cli --catalog delta
```

Delta Lake 테이블에 쓰기

다음 SQL 명령을 사용하여 테이블을 생성하고 테이블에 씁니다.

```
SHOW SCHEMAS;

CREATE TABLE default.delta_table (id  int, data varchar, category varchar) WITH
( location =  's3://amzn-s3-demo-bucket/<prefix>');

INSERT INTO default.delta_table VALUES  (1,'a','c1'), (2,'b','c2'), (3,'c','c3');
```

Delta Lake 테이블에서 읽기

다음 SQL 명령을 사용하여 테이블에서 읽습니다.

```

SELECT * from default.delta_table;
```

Spark와 함께 Delta Lake 클러스터 사용

Amazon EMR 버전 6.9.0부터 부트스트랩 작업을 포함할 필요 없이 Spark 클러스터에서 Delta Lake를 사용할 수 있습니다. Amazon EMR 버전 6.8.0 이하의 경우 부트스트랩 작업을 사용하여 필요한 모든 종속 항목을 사전 설치할 수 있습니다.

다음 예제에서는 AWS CLI 를 사용하여 Amazon EMR Spark 클러스터에서 Delta Lake로 작업합니다.

에서 Amazon EMR의 Delta Lake를 사용하려면 AWS Command Line Interface먼저 클러스터를 생성합니다. Delta Lake 분류를 지정하는 방법에 대한 자세한 내용은 [클러스터를 생성할 AWS Command Line Interface 때를 사용하여 구성 제공] 또는 클러스터를 생성할 때 Java SDK로 구성 제공을 AWS Command Line Interface참조하세요. https://docs.aws.amazon.com/emr/latest/ReleaseGuide/emr-configure-apps-create-cluster.html#emr-configure-apps-create-cluster-sdk

1. 다음 콘텐츠가 포함된 `configurations.json` 파일을 생성합니다.

```
[{"Classification":"delta-defaults",  "Properties":{"delta.enabled":"true"} }]
```

2. 다음과 같은 구성으로 클러스터를 생성하고 Amazon S3 **bucket path** 및 **subnet ID** 예제를 사용자 정보로 바꿉니다.

```
aws emr create-cluster
```

```
    --release-label  emr-6.9.0
    --applications Name=Spark
    --configurations file://delta_configurations.json
    --region us-east-1
    --name My_Spark_Delta_Cluster
    --log-uri  s3://amzn-s3-demo-bucket/
    --instance-type m5.xlarge
    --instance-count 2
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

또는 Spark 작업에서 다음 파일을 JAR 종속 항목으로 사용하여 Amazon EMR 클러스터와 Spark 애플리케이션을 생성할 수 있습니다.

```
/usr/share/aws/delta/lib/delta-core.jar,
/usr/share/aws/delta/lib/delta-storage.jar,
/usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar
```

> **Note**
>
> Amazon EMR 릴리스 7.0.0 이상을 사용하는 경우 `/usr/share/aws/delta/lib/delta-spark.jar` 대신를 사용합니다`/usr/share/aws/delta/lib/delta-core.jar`.

자세한 내용은 [애플리케이션 제출]을 참조하세요.

jar를 Spark 작업에 jar 종속 항목으로 포함하려면 Spark 애플리케이션에 다음 구성 속성을 추가할 수 있습니다.

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-core.jar,
    /usr/share/aws/delta/lib/delta-storage.jar,
    /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Spark 작업 종속 항목에 대한 자세한 내용은 [Dependency Management]를 참조하세요.

Amazon EMR 릴리스 7.0.0 이상을 사용하는 경우 대신 `/usr/share/aws/delta/lib/delta-spark.jar` 구성을 추가합니다.

```
--conf “spark.jars=/usr/share/aws/delta/lib/delta-spark.jar,
     /usr/share/aws/delta/lib/delta-storage.jar,
     /usr/share/aws/delta/lib/delta-storage-s3-dynamodb.jar"
```

Delta Lake용 Spark 세션 초기화

다음 예제에서는 대화식 Spark 셸을 시작하거나 Spark 제출을 사용하거나 Amazon EMR에서 Delta Lake를 작업하기 위해 Amazon EMR Notebooks를 사용하는 방법을 보여줍니다.

spark-shell

1. SSH를 사용하여 프라이머리 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. Spark 셸을 시작하려면 다음 명령을 입력합니다. PySpark 셸을 사용하려면 `spark-shell`을 `pyspark`로 바꿉니다.

```
spark-shell \
   --conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension" \
   --conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

   Amazon EMR 릴리스 6.15.0 이상을 실행하는 경우 다음 구성을 사용하여 Delta Lake에서 Lake Formation을 기반으로 세분화된 액세스 제어를 사용해야 합니다.

```
spark-shell \
  --conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
 \
  --conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

  --conf spark.sql.catalog.spark_catalog.lf.managed=true
```

spark-submit

1. SSH를 사용하여 프라이머리 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.

2. Delta Lake용 Spark 세션을 시작하려면 다음 명령을 입력합니다.

```
spark-submit
—conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension"
—conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"
```

Amazon EMR 릴리스 6.15.0 이상을 실행하는 경우 다음 구성을 사용하여 Delta Lake에서 Lake Formation을 기반으로 세분화된 액세스 제어를 사용해야 합니다.

```
spark-submit \  `
--conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserv
--conf
 spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog \

--conf spark.sql.catalog.spark_catalog.lf.managed=true
```

EMR Studio notebooks

Amazon EMR Studio 노트북을 사용하여 Spark 세션을 초기화하려면 다음 예제와 같이 Amazon EMR Notebooks에서 %%configure 매직 명령을 사용하여 Spark 세션을 구성합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR Notebooks 매직 사용을 참조하세요.

```
%%configure -f
{
  "conf": {
    "spark.sql.extensions":  "io.delta.sql.DeltaSparkSessionExtension",
     "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog"
  }
}
```

Amazon EMR 릴리스 6.15.0 이상을 실행하는 경우 다음 구성을 사용하여 Delta Lake에서 Lake Formation을 기반으로 세분화된 액세스 제어를 사용해야 합니다.

```
%%configure -f
{
  "conf": {
```

```
    "spark.sql.extensions":
 "io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserver.connector.spark.sql
    "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.delta.catalog.DeltaCatalog",
    "spark.sql.catalog.spark_catalog.lf.managed": "true"
  }
}
```

Delta Lake 테이블에 쓰기

다음 예제에서는 Spark DataFrame을 생성하고 Delta Lake 데이터 세트로 쓰는 방법을 보여줍니다. 이 예제에서는 SSH를 기본 Hadoop 사용자로 사용하여 프라이머리 노드에 연결된 상태에서 Spark 셸로 데이터 세트를 처리하는 방법을 보여줍니다.

> **Note**
>
> 코드 샘플을 Spark 셸에 붙여넣으려면 프롬프트에 :paste를 입력하고 예제를 붙여넣은 다음 CTRL + D를 누릅니다.

PySpark

Spark에는 Python 기반 셸인 `pyspark`도 포함되어 있으며, 이 셸을 사용하여 Python에서 작성된 Spark 프로그램을 시제품화할 수 있습니다. `spark-shell`과 마찬가지로 프라이머리 노드에서 `pyspark`를 간접 호출합니다.

```
## Create a DataFrame
data =  spark.createDataFrame([("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")],
["id", "creation_date",  "last_update_time"])

## Write a DataFrame as a Delta Lake dataset to the S3  location
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string, creation_date
 string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");
```

```
data.writeTo("delta_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame
val data = Seq(("100",  "2015-01-01",  "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01",  "2015-01-01T13:51:40.519832Z")).toDF("id", "creation_date",
 "last_update_time")

// Write a DataFrame as a Delta Lake dataset to the S3  location
spark.sql("""CREATE  TABLE IF NOT EXISTS delta_table (id string,
creation_date string,
last_update_time string)
USING delta location
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table'""");

data.write.format("delta").mode("append").saveAsTable("delta_table")
```

SQL

```
-- Create a Delta  Lake table with the S3 location
CREATE TABLE delta_table(id string,
creation_date string,
last_update_time string)
USING delta LOCATION
's3://amzn-s3-demo-bucket/example-prefix/db/delta_table';

-- insert data into the table
INSERT INTO delta_table VALUES  ("100", "2015-01-01",
 "2015-01-01T13:51:39.340396Z"),
("101",  "2015-01-01",  "2015-01-01T12:14:58.597216Z"),
("102",  "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103",  "2015-01-01", "2015-01-01T13:51:40.519832Z");
```

Delta Lake 테이블에서 읽기

PySpark

```
ddf = spark.table("delta_table")
ddf.show()
```

Scala

```
val ddf =  spark.table("delta_table")
ddf.show()
```

SQL

```
SELECT * FROM delta_table;
```

Spark 및 AWS Glue와 함께 Delta Lake 클러스터 사용

AWS Glue 카탈로그를 Delta Lake 테이블의 메타스토어로 사용하려면 다음 단계를 수행하여 클러스터를 생성합니다. 를 사용하여 Delta Lake 분류를 지정하는 방법에 대한 자세한 내용은 [클러스터를 생성할 AWS Command Line Interface 때를 사용하여 구성 제공](#) 또는 클러스터를 생성할 때 Java SDK를 사용하여 구성 제공을 AWS Command Line Interface참조하세요. https://docs.aws.amazon.com/emr/latest/ReleaseGuide/emr-configure-apps-create-cluster.html#emr-configure-apps-create-cluster-sdk

Delta Lake 클러스터 생성

1. 다음 콘텐츠가 포함된 `configurations.json` 파일을 생성합니다.

```
[{"Classification":"delta-defaults",
"Properties":{"delta.enabled":"true"}},
{"Classification":"spark-hive-site",
"Properties":
{"hive.metastore.client.factory.class":"com.amazonaws.glue.catalog.metastore.AWSGlueDataCat
```

2. 다음과 같은 구성으로 클러스터를 생성하고 **example Amazon S3 bucket path** 및 **subnet ID**를 사용자 정보로 바꿉니다.

```
aws emr create-cluster
    --release-label  emr-6.9.0
    --applications Name=Spark
    --configurations file://delta_configurations.json
    --region us-east-1
    --name My_Spark_Delta_Cluster
    --log-uri  s3://amzn-s3-demo-bucket/
    --instance-type m5.xlarge
    --instance-count 2
    --service-role EMR_DefaultRole_V2
    --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

고려 사항 및 제한 사항

- Delta Lake는 Amazon EMR 릴리스 6.9.0 이상에서 함께 사용할 수 있습니다. Delta 테이블이 있는 Amazon EMR 클러스터에서 Apache Spark 3.x를 사용할 수 있습니다.
- 최상의 성능, 보안 및 안정성을 위해 s3a 대신 S3 위치 경로에 s3 URI 스키마를 사용하는 것이 좋습니다. 자세한 내용은 스토리지 및 파일 시스템 작업을 참조하세요.
- Amazon EMR 7.0에서는 Delta Universal Format (UniForm) 및 convert-to-Iceberg 명령문이 지원되지 않습니다.
- Amazon EMR 6.9 및 6.10에서 Amazon S3에 Delta Lake 테이블 데이터를 저장하면 열 이름 바꾸기 작업 후에 열 데이터가 `NULL`이 됩니다. 이 문제는 Amazon EMR 6.11에서 해결되었습니다. 이 실험적인 열 이름 바꾸기 작업에 대한 자세한 내용은 Delta Lake 사용 설명서에서 Column rename operation을 참조하세요.
- Apache Spark 외부에서 AWS Glue 데이터 카탈로그에 데이터베이스를 생성하는 경우 데이터베이스에 빈 `LOCATION` 필드가 있을 수 있습니다. Spark에서는 빈 위치 속성으로 데이터베이스를 생성할 수 없으므로 Amazon EMR에서 Spark를 사용하여 Glue 데이터베이스에 Delta 테이블을 생성하고 데이터베이스에 빈 속성이 있는 경우 `LOCATION` 오류가 발생합니다.

  ```
  IllegalArgumentException: Can not create a Path from an empty string
  ```

  이 문제를 해결하려면 `LOCATION` 필드에 대한 비어 있지 않은 유효한 경로를 사용하여 데이터 카탈로그에서 데이터베이스를 생성합니다. 이 솔루션을 구현하는 단계는 Amazon Athena 사용 설명서에서 테이블을 생성할 때 잘못된 인수 예외 발생을 참조하세요.

Delta 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Delta의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 [Amazon EMR 7.x 릴리스 버전](#), [Amazon EMR 6.x 릴리스 버전](#) 또는 [Amazon EMR 5.x 릴리스 버전](#)의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Delta 버전 정보

Amazon EMR 릴리스 레이블	Delta 버전	Delta와 함께 설치된 구성 요소
emr-7.9.0	3.3.0-amzn-1	Not available.
emr-7.8.0	3.3.0-amzn-0	Not available.
emr-7.7.0	3.2.1-amzn-2	Not available.
emr-7.6.0	3.2.1-amzn-1	Not available.
emr-7.5.0	3.2.1-amzn-0	Not available.
emr-7.4.0	3.2.0-amzn-1	Not available.
emr-7.3.0	3.2.0-amzn-0	Not available.
emr-7.2.0	3.1.0	Not available.
emr-7.1.0	3.0.0	Not available.
emr-7.0.0	3.0.0	Not available.
emr-6.15.0	2.4.0	Not available.
emr-6.14.0	2.4.0	Not available.
emr-6.13.0	2.4.0	Not available.
emr-6.12.0	2.4.0	Not available.
emr-6.11.1	2.2.0	Not available.
emr-6.11.0	2.2.0	Not available.

Amazon EMR 릴리스 레이블	Delta 버전	Delta와 함께 설치된 구성 요소
emr-6.10.1	2.2.0	Not available.
emr-6.10.0	2.2.0	Not available.
emr-6.9.1	2.1.0	Not available.
emr-6.9.0	2.1.0	Not available.

Apache Flink

Apache Flink는 처리량이 많은 데이터 원본에서 실시간 스트림 처리를 실행하는 데 사용할 수 있는 스트리밍 데이터플로우 엔진입니다. Flink는 잘못된 순서 이벤트에 대한 이벤트 시간 의미, 일회성 의미, 역압 제어, 그리고 스트리밍과 배치 애플리케이션 모두를 작성하도록 최적화된 API를 지원합니다.

또한, Flink에는 다음과 같은 타사 데이터 원본용 커넥터가 있습니다.

- Amazon Kinesis Data Streams
- Apache Kafka
- Flink Elasticsearch 커넥터
- Twitter Streaming API
- Cassandra

Amazon EMR에서는 클러스터에 있는 다른 애플리케이션과 함께 리소스를 관리할 수 있도록 YARN 애플리케이션으로 Flink를 지원합니다. Flink-on-YARN을 사용하여 임시 Flink 작업을 제출하거나 여러 작업을 허용하고 전체 YARN 동일 시작 인스턴스에 따라 리소스를 할당하는 장기 실행 클러스터를 생성할 수 있습니다.

Flink는 Amazon EMR 릴리스 버전 5.1.0 이상에 포함되어 있습니다.

> **Note**
>
> `FlinkKinesisConsumer` 클래스에 대한 지원이 Amazon EMR 릴리스 버전 5.2.1에 추가되었습니다.

다음 테이블에는 Amazon EMR이 Flink를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Flink의 버전이 나열되어 있습니다.

이 릴리스에서 Flink와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Flink 버전 정보

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-7.9.0	Flink 1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datano

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
		de, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

다음 테이블에는 Amazon EMR이 Flink를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Flink의 버전이 나열되어 있습니다.

이 릴리스에서 Flink와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Flink 버전 정보

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-6.15.0	Flink 1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

다음 테이블에는 Amazon EMR이 Flink를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Flink의 버전이 나열되어 있습니다.

이 릴리스에서 Flink와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Flink 버전 정보

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.36.2	Flink 1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

주제

Flink를 포함하는 클러스터 생성

AWS Management Console AWS CLI또는 AWS SDK를 사용하여 클러스터를 시작할 수 있습니다.

콘솔에서 Flink가 설치된 클러스터를 시작하는 방법

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.

3. 소프트웨어 구성에서 EMR 릴리스 emr-5.1.0 이상을 선택합니다.
4. 설치할 다른 애플리케이션과 함께 Flink를 애플리케이션으로 선택합니다.
5. 필요에 따라 다른 옵션을 선택하고 클러스터 생성을 선택합니다.

에서 Flink를 사용하여 클러스터를 시작하려면 AWS CLI

- 다음 명령을 사용하여 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=MyKeyName,InstanceProfile=EMR_EC2_DefaultRole \
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args=flink-yarn-session,-d
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

Amazon EMR에서 Flink 구성

Hive 메타스토어 및 Glue 카탈로그를 사용하여 Flink 구성

Amazon EMR 릴리스 6.9.0 이상은 Hive에 대한 Apache Flink 커넥터를 사용하여 Hive 메타스토어와 AWS Glue 카탈로그를 모두 지원합니다. 이 섹션에서는 Flink를 사용하여 AWS Glue 카탈로그 및 Hive 메타스토어를 구성하는 데 필요한 단계를 설명합니다.

주제

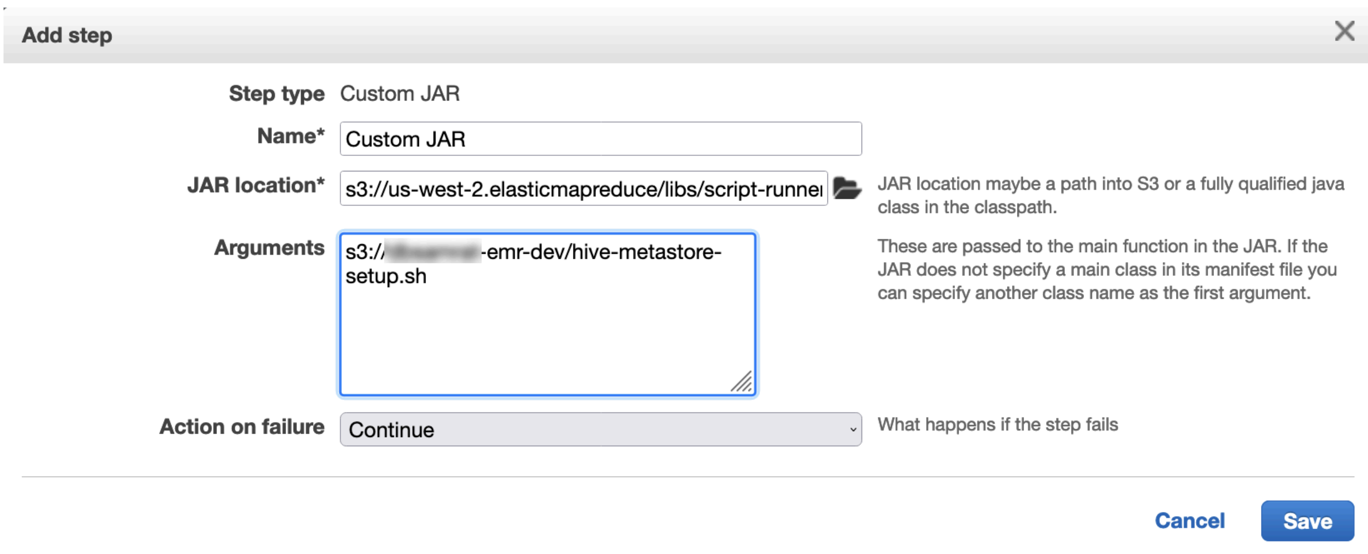
- Hive 메타스토어 사용
- AWS Glue 데이터 카탈로그 사용

Hive 메타스토어 사용

1. 릴리스 6.9.0 이상 및 둘 이상의 애플리케이션(Hive 및 Flink)이 포함된 EMR 클러스터를 생성합니다.
2. 스크립트 러너를 사용하여 다음 스크립트를 단계 함수로 실행합니다.

   hive-metastore-setup.sh

```
sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
```



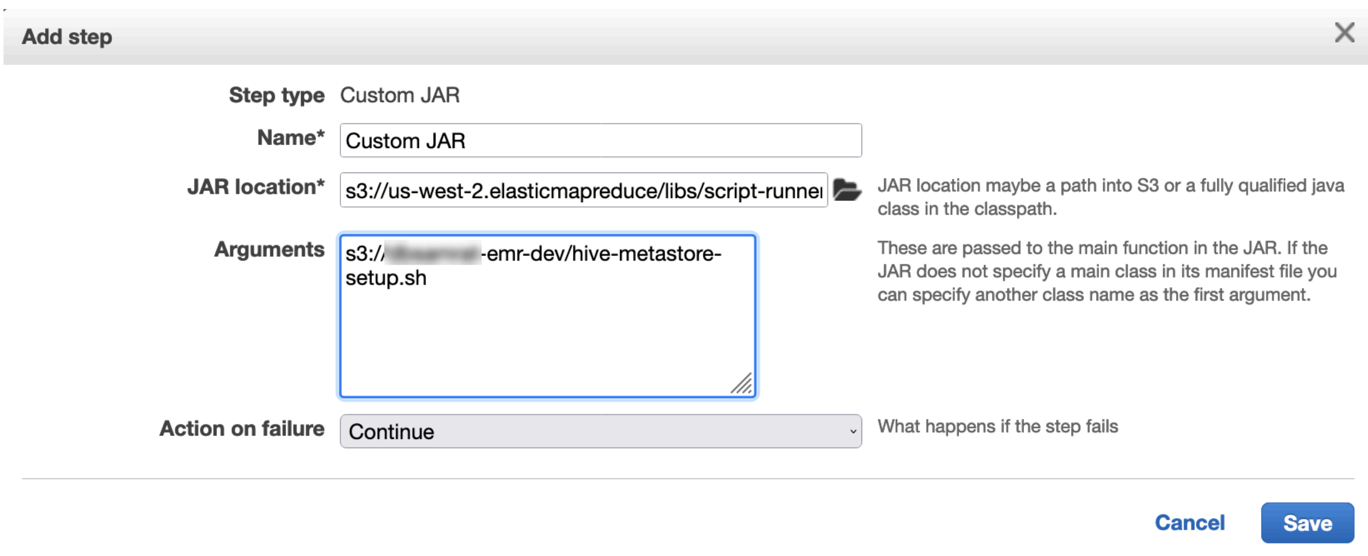
AWS Glue 데이터 카탈로그 사용

1. 릴리스 6.9.0 이상 및 둘 이상의 애플리케이션(Hive 및 Flink)이 포함된 EMR 클러스터를 생성합니다.
2. AWS Glue 데이터 카탈로그 설정에서 Hive 테이블 메타데이터에 사용을 선택하여 클러스터에서 데이터 카탈로그를 활성화합니다.

3. [스크립트 러너](#)를 사용하여 다음 스크립트를 단계 함수로 실행합니다([Amazon EMR 클러스터에서 명령 및 스크립트 실행](#)).

   glue-catalog-setup.sh

```
sudo cp /usr/lib/hive/auxlib/aws-glue-datacatalog-hive3-client.jar /usr/lib/flink/
lib
sudo cp /usr/lib/hive/lib/antlr-runtime-3.5.2.jar /usr/lib/flink/lib
sudo cp /usr/lib/hive/lib/hive-exec-3.1.3*.jar /lib/flink/lib
sudo cp /usr/lib/hive/lib/libfb303-0.9.3.jar /lib/flink/lib
sudo cp /usr/lib/flink/opt/flink-connector-hive_2.12-1.15.2.jar /lib/flink/lib
sudo chmod 755 /usr/lib/flink/lib/aws-glue-datacatalog-hive3-client.jar
sudo chmod 755 /usr/lib/flink/lib/antlr-runtime-3.5.2.jar
sudo chmod 755 /usr/lib/flink/lib/hive-exec-3.1.3*.jar
sudo chmod 755 /usr/lib/flink/lib/libfb303-0.9.3.jar
sudo chmod 755 /usr/lib/flink/lib/flink-connector-hive_2.12-1.15.2.jar
```



구성 파일을 사용하여 Flink 구성

Amazon EMR 구성 API를 사용하여 구성 파일로 Flink를 구성할 수 있습니다. 현재 API 내에서 구성할 수 있는 파일은 다음과 같습니다.

- `flink-conf.yaml`
- `log4j.properties`
- `flink-log4j-session`

- `log4j-cli.properties`

Flink용 기본 구성 파일은 `flink-conf.yaml`입니다.

AWS CLI를 사용하여 Flink에 사용되는 작업 슬롯 수를 구성하는 방법

1. 다음 콘텐츠가 포함된 `configurations.json` 파일을 생성합니다.

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "taskmanager.numberOfTaskSlots":"2"
      }
    }
]
```

2. 다음 구성을 사용하여 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

> **Note**
>
> Flink API를 사용하여 일부 구성을 변경할 수도 있습니다. 자세한 내용은 Flink 설명서에서 [Concepts]를 참조하세요.
> Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 [실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공]을 참조하십시오.

병렬 처리 옵션

애플리케이션 소유자는 Flink 내에서 작업에 할당할 리소스를 잘 알고 있습니다. 이 설명서의 예제에서는 애플리케이션에 사용되는 태스크 인스턴스와 동일한 개수의 작업을 사용합니다. 일반적으로 이러한 사용은 초기 수준의 병렬화에 권장되지만, 작업 슬롯을 사용하여 더 세부적으로 병렬화할 수도 있습니다. 이 경우 일반적으로 인스턴스당 가상 코어 수를 초과하면 안 됩니다. Flink의 아키텍처에 대한 자세한 내용은 Flink 설명서에서 Concepts를 참조하세요.

여러 프라이머리 노드를 포함하는 EMR 클러스터에서 Flink 구성

Flink의 JobManager는 여러 프라이머리 노드가 있는 Amazon EMR 클러스터에서 프라이머리 노드 장애 조치 프로세스 중에도 계속 사용할 수 있습니다. Amazon EMR 5.28.0부터 JobManager 고가용성이 자동으로 활성화됩니다. 수동 구성이 필요하지 않습니다.

Amazon EMR 버전 5.27.0 또는 이하 버전에서 JobManager는 단일 장애 지점입니다. JobManager가 실패하면 모든 작업 상태가 손실되고 실행 중인 작업이 다시 시작되지 않습니다. 다음 예제에서 보여주듯이, 애플리케이션 시도 횟수를 구성하고 체크포인트를 수행하고 Flink의 상태 스토리지로 ZooKeeper를 활성화하여 JobManager 고가용성을 활성화할 수 있습니다.

```
[
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.resourcemanager.am.max-attempts": "10"
    }
  },
  {
    "Classification": "flink-conf",
    "Properties": {
        "yarn.application-attempts": "10",
        "high-availability": "zookeeper",
        "high-availability.zookeeper.quorum": "%{hiera('hadoop::zk')}",
        "high-availability.storageDir": "hdfs:///user/flink/recovery",
        "high-availability.zookeeper.path.root": "/flink"
    }
  }
]
```

YARN에 대한 최대 애플리케이션 마스터 시도와 Flink에 대한 애플리케이션 시도를 모두 구성해야 합니다. 자세한 내용은 Configuration of YARN cluster high availability를 참조하세요. 또한 다시 시작된

JobManager가 이전에 완료된 체크포인트에서 실행 중인 작업을 복원하도록 Flink 체크포인트를 구성할 수 있습니다. 자세한 내용은 Flink checkpointing을 참조하세요.

메모리 프로세스 크기 구성

Flink 1.11.x를 사용하는 Amazon EMR 버전의 경우 `flink-conf.yaml`에서 JobManager(`jobmanager.memory.process.size`) 및 TaskManager(`taskmanager.memory.process.size`) 모두에 대해 총 메모리 프로세스 크기를 구성해야 합니다. 구성 API로 클러스터를 구성하거나 SSH를 통해 이러한 필드의 주석 처리를 수동으로 제거하여 이러한 값을 설정할 수 있습니다. Flink는 다음과 같은 기본값을 제공합니다.

- `jobmanager.memory.process.size`: 1,600m
- `taskmanager.memory.process.size`: 1,728m

JVM 메타스페이스와 오버헤드를 제외하려면 `taskmanager.memory.process.size` 대신 Flink 총 메모리 크기(`taskmanager.memory.flink.size`)를 대신 사용합니다. `taskmanager.memory.process.size`의 기본값은 1,280m입니다. `taskmanager.memory.process.size` 및 `taskmanager.memory.process.size`를 모두 설정하는 것은 권장되지 않습니다.

Flink 1.12.0 이상을 사용하는 모든 Amazon EMR 버전에는 Flink용 오픈 소스 세트에 Amazon EMR의 기본값으로 기본값이 나열되어 있으므로 사용자가 직접 구성할 필요가 없습니다.

로그 출력 파일 크기 구성

Flink 애플리케이션 컨테이너는 `.out` 파일, `.log` 파일, `.err` 파일과 같은 세 가지 유형의 로그 파일을 생성하고 작성합니다. `.err` 파일만 압축되어 파일 시스템에서 제거되고, `.log` 및 `.out` 로그 파일은 파일 시스템에 남습니다. 이러한 출력 파일을 관리할 수 있고 클러스터를 안정적으로 유지하기 위해 최대 파일 수를 설정하고 파일 크기를 제한하도록 `log4j.properties`에서 로그 로테이션을 구성할 수 있습니다.

Amazon EMR 버전 5.30.0 이상

Amazon EMR 5.30.0부터 Flink는 구성 분류 이름 `flink-log4j.`의 log4j2 로깅 프레임워크를 사용합니다. 다음 예제 구성은 log4j2 형식을 보여줍니다.

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
```

```
      "appender.main.name": "MainAppender",
      "appender.main.type": "RollingFile",
      "appender.main.append" : "false",
      "appender.main.fileName" : "${sys:log.file}",
      "appender.main.filePattern" : "${sys:log.file}.%i",
      "appender.main.layout.type" : "PatternLayout",
      "appender.main.layout.pattern" : "%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p %-60c %x - %m
%n",
      "appender.main.policies.type" : "Policies",
      "appender.main.policies.size.type" : "SizeBasedTriggeringPolicy",
      "appender.main.policies.size.size" : "100MB",
      "appender.main.strategy.type" : "DefaultRolloverStrategy",
      "appender.main.strategy.max" : "10"
    },
  }
]
```

Amazon EMR 버전 5.29.0 이하

Amazon EMR 버전 5.29.0 이하에서 Flink는 log4j 로깅 프레임워크를 사용합니다. 다음 구성 예제에서는 log4j 형식을 보여줍니다.

```
[
  {
    "Classification": "flink-log4j",
    "Properties": {
      "log4j.appender.file": "org.apache.log4j.RollingFileAppender",
      "log4j.appender.file.append":"true",
      # keep up to 4 files and each file size is limited to 100MB
      "log4j.appender.file.MaxFileSize":"100MB",
      "log4j.appender.file.MaxBackupIndex":4,
      "log4j.appender.file.layout":"org.apache.log4j.PatternLayout",
      "log4j.appender.file.layout.ConversionPattern":"%d{yyyy-MM-dd HH:mm:ss,SSS} %-5p
 %-60c %x - %m%n"
    },
  }
]
```

Java 11과 함께 실행하도록 Flink 구성

Amazon EMR 릴리스 6.12.0 이상에서는 Flink에 대해 Java 11 런타임 지원을 제공합니다. 다음 섹션에서는 Flink에 대한 Java 11 런타임 지원을 제공하도록 클러스터를 구성하는 방법을 설명합니다.

주제

- [클러스터를 생성할 때 Java 11용 Flink 구성](#)
- [실행 중인 클러스터에서 Java 11용 Flink 구성](#)
- [실행 중인 클러스터에서 Flink의 Java 런타임 확인](#)

클러스터를 생성할 때 Java 11용 Flink 구성

다음 단계에 따라 Flink 및 Java 11 런타임으로 EMR 클러스터를 생성합니다. Java 11 런타임 지원을 추가하는 구성 파일은 `flink-conf.yaml`입니다.

Console

콘솔에서 Flink 및 Java 11 런타임이 포함된 클러스터를 생성하는 방법

1. 에 로그인 AWS Management Console하고 https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 탐색 창의 EMR on EC2에서 클러스터를 선택하고 클러스터 생성을 선택합니다.
3. Amazon EMR 릴리스 6.12.0 이상을 선택하고 Flink 애플리케이션을 설치하도록 선택합니다. 클러스터에 설치하려는 다른 애플리케이션을 모두 선택합니다.
4. 클러스터를 계속 설정합니다. 선택적 소프트웨어 설정 섹션에서 기본 구성 입력 옵션을 사용하고 다음 구성을 입력합니다.

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
]
```

5. 계속 설정하고 클러스터를 시작합니다.

AWS CLI

CLI에서 Flink 및 Java 11 런타임이 포함된 클러스터를 생성하는 방법

1. Java 11을 사용하도록 Flink를 구성하는 구성 파일 `configurations.json`을 생성합니다.

```
[
    {
      "Classification": "flink-conf",
      "Properties": {
        "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
        "env.java.home":"/usr/lib/jvm/jre-11"
      }
    }
]
```

2. 다음 예제와 같이에서 Amazon EMR 릴리스 6.12.0 이상을 사용하여 AWS CLI새 EMR 클러스터를 생성하고 Flink 애플리케이션을 설치합니다.

```
aws emr create-cluster --release-label emr-6.12.0 \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=YourKeyName,InstanceProfile=EMR_EC2_DefaultRole
```

실행 중인 클러스터에서 Java 11용 Flink 구성

다음 단계를 사용하여 Flink 및 Java 11 런타임으로 실행 중인 EMR 클러스터를 업데이트합니다. Java 11 런타임 지원을 추가하는 구성 파일은 `flink-conf.yaml`입니다.

Console

콘솔에서 Flink 및 Java 11 런타임으로 실행 중인 클러스터를 업데이트하는 방법

1. 에 로그인 AWS Management Console하고 https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.

2. 탐색 창에서 EMR on EC2 아래 클러스터를 선택하고 업데이트할 클러스터를 선택합니다.

   Note

   클러스터가 Java 11을 지원하려면 Amazon EMR 릴리스 6.12.0 이상을 사용해야 합니다.

3. 구성 탭을 선택합니다.
4. 인스턴스 그룹 구성 섹션에서 업데이트하려는 실행 중인 인스턴스 그룹을 선택한 다음, 목록 작업 메뉴에서 재구성을 선택합니다.
5. 다음과 같이 속성 편집 옵션을 사용하여 인스턴스 그룹을 재구성합니다. 각 항목 이후에 새 구성 추가를 선택합니다.

분류	속성	값
`flink-conf`	`containerized.taskmanager.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`containerized.master.env.JAVA_HOME`	`/usr/lib/jvm/jre-11`
`flink-conf`	`env.java.home`	`/usr/lib/jvm/jre-11`

6. 변경 내용 저장을 선택하여 구성을 추가합니다.

AWS CLI

CLI에서 Flink 및 Java 11 런타임을 사용하도록 실행 중인 클러스터를 업데이트하는 방법

`modify-instance-groups` 명령을 사용하여 실행 중인 클러스터의 인스턴스 그룹에 대해 새 구성을 지정합니다.

1. 먼저 Java 11을 사용하도록 Flink를 구성하는 `configurations.json` 구성 파일을 생성합니다. 다음 예제에서는 *ig-1xxxxxxx9*를 재구성하려는 인스턴스 그룹의 ID로 바꿉니다. `modify-instance-groups` 명령을 실행하는 동일한 디렉터리에 파일을 저장합니다.

```
[
   {
      "InstanceGroupId":"ig-1xxxxxxx9",
      "Configurations":[
         {
            "Classification":"flink-conf",
            "Properties":{
              "containerized.taskmanager.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "containerized.master.env.JAVA_HOME":"/usr/lib/jvm/jre-11",
              "env.java.home":"/usr/lib/jvm/jre-11"
            },
            "Configurations":[]
         }
      ]
   }
]
```

2. 에서 다음 명령을 AWS CLI실행합니다. 재구성하려는 인스턴스 그룹의 ID를 바꿉니다.

```
aws emr modify-instance-groups --cluster-id j-2AL4XXXXXX5T9 \
--instance-groups file://configurations.json
```

실행 중인 클러스터에서 Flink의 Java 런타임 확인

실행 중인 클러스터에 대한 Java 런타임을 확인하려면 SSH를 사용하여 프라이머리 노드에 연결에서 설명한 대로 SSH를 사용하여 프라이머리 노드에 로그인합니다. 그런 다음, 다음 명령을 실행합니다.

```
ps -ef | grep flink
```

`-ef` 옵션과 함께 ps 명령은 시스템에서 실행 중인 모든 프로세스를 나열합니다. `grep`로 해당 출력을 필터링하여 `flink` 문자열에 대한 언급을 찾을 수 있습니다. Java 런타임 환경(JRE) 값(`jre-XX`)에 대한 출력을 검토합니다. 다음 출력에서 `jre-11`은 런타임에 Flink용으로 Java 11이 선택되었음을 나타냅니다.

```
flink    19130     1  0 09:17 ?        00:00:15 /usr/lib/jvm/jre-11/bin/
java -Djava.io.tmpdir=/mnt/tmp -Dlog.file=/usr/lib/flink/log/flink-flink-
historyserver-0-ip-172-31-32-127.log -Dlog4j.configuration=file:/usr/lib/flink/conf/
log4j.properties -Dlog4j.configurationFile=file:/usr/lib/flink/conf/log4j.properties
 -Dlogback.configurationFile=file:/usr/lib/flink/conf/logback.xml -classpath /usr/lib/
flink/lib/flink-cep-1.17.0.jar:/usr/lib/flink/lib/flink-connector-files-1.17.0.jar:/
```

```
usr/lib/flink/lib/flink-csv-1.17.0.jar:/usr/lib/flink/lib/flink-json-1.17.0.jar:/usr/
lib/flink/lib/flink-scala_2.12-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-java-
uber-1.17.0.jar:/usr/lib/flink/lib/flink-table-api-scala-bridge_2.12-1.17.0.
```

또는 SSH를 사용하여 프라이머리 노드에 로그인하고 `flink-yarn-session -d` 명령을 사용하여 Flink YARN 세션을 시작합니다. 출력은 다음 예제에서 Flink용 가상 머신(JVM)(`java-11-amazon-corretto`)을 보여줍니다.

```
2023-05-29 10:38:14,129 INFO  org.apache.flink.configuration.GlobalConfiguration
     [] - Loading configuration property: containerized.master.env.JAVA_HOME, /usr/lib/
jvm/java-11-amazon-corretto.x86_64
```

Amazon EMR에서 Flink 작업 사용

콘솔을 통해 Amazon EMR 기반 Flink와 상호 작용하는 몇 가지 방법이 있으며, Flink 인터페이스는 ResourceManager Tracking UI와 명령줄에 있습니다. 이들 중 하나를 사용하여 JAR 파일을 Flink 애플리케이션에 제출할 수 있습니다. JAR 파일을 제출하면 Flink JobManager에서 관리하는 작업이 됩니다. JobManager는 Flink 세션 애플리케이션 마스터 대몬(daemon)을 호스팅하는 YARN 노드에 있습니다.

또한 Flink 애플리케이션을 장기 실행 YARN 작업이나 임시 클러스터로 실행할 수 있습니다. 장기 실행 클러스터에서는 Amazon EMR 클러스터에서 실행 중인 Flink 클러스터 하나로 여러 Flink 애플리케이션을 제출할 수 있습니다. Flink 작업을 임시 클러스터에서 실행할 경우 Amazon EMR 클러스터는 Flink 애플리케이션을 실행하는 데 걸리는 시간 동안만 존재하므로 사용한 리소스와 시간에 대해서만 요금이 청구됩니다. Amazon EMR `AddSteps` API 작업, 작업에 대한 단계 인수, 또는 `create-cluster` 명령을 `RunJobFlow` 통해 Flink 작업을 제출할 수 있습니다 AWS CLI `add-steps`.

장기 실행 클러스터에서 한 단계로 Flink YARN 애플리케이션 시작

여러 클라이언트가 YARN API 작업을 통해 작업을 제출할 수 있는 Flink 애플리케이션을 시작하려면 클러스터를 생성하거나 Flink 애플리케이션을 기존 클러스터에 추가해야 합니다. 새 클러스터를 생성하는 방법에 대한 지침은 Flink를 포함하는 클러스터 생성 부분을 참조하세요. 기존 클러스터에서 YARN 세션을 시작하려면 콘솔, AWS CLI또는 Java SDK에서 다음 단계를 수행합니다.

> **Note**
>
> `flink-yarn-session` 명령은 실행을 간소화하기 위한 `yarn-session.sh` 스크립트 래퍼로 Amazon EMR 버전 5.5.0에 추가되었습니다. 이전 버전의 Amazon EMR을 사용하는 경우

콘솔에서 인수에 대해 `bash -c "/usr/lib/flink/bin/yarn-session.sh -d"`를 대체하거나 AWS CLI 명령에서 `Args`를 대체합니다.

콘솔에서 기존 클러스터의 Flink 작업을 제출하는 방법

기존 클러스터에서 `flink-yarn-session` 명령을 사용하여 Flink 세션을 제출합니다.

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 목록에서 이전에 시작한 클러스터를 선택합니다.
3. 클러스터 세부 정보 페이지에서 단계, Add Step(단계 추가)을 선택합니다.
4. 아래 지침에 따라 파라미터를 입력하고 추가를 선택합니다.

파라미터	설명
단계 유형	사용자 지정 JAR
이름	단계 식별을 위한 이름. 예를 들어, *<example-flink-step-name>* 과 같습니다.
Jar location(Jar 위치)	command-runner.jar
인수	애플리케이션에 따라 여러 가지 인수를 사용하는 `flink-yarn-session` 명령. 예를 들어 *`flink-yarn-session -d`* 는 YARN 클러스터 내에서 분리된 상태(-d)로 Flink 세션을 시작합니다. 인수에 대한 자세한 내용은 최신 Flink 설명서에서 YARN Setup을 참조하세요.

를 사용하여 기존 클러스터에서 Flink 작업을 제출하려면 AWS CLI

- `add-steps` 명령을 사용하여 장기 실행 클러스터에 Flink 작업을 추가합니다. 다음 예제 명령은 YARN 클러스터 내에서 분리된 상태(`-d`)로 Flink 세션을 시작하도록 `Args="flink-yarn-session", "-d"`를 지정합니다. 인수에 대한 자세한 내용은 최신 Flink 설명서에서 YARN Setup을 참조하세요.

```
aws emr add-steps --cluster-id <j-XXXXXXXX> --steps Type=CUSTOM_JAR,Name=<example-
flink-step-name>,Jar=command-runner.jar,Args="flink-yarn-session","-d"
```

장기 실행 클러스터에서 기존 Flink 애플리케이션에 작업 제출

장기 실행 클러스터에 기존 Flink 애플리케이션이 이미 있는 경우 클러스터의 Flink 애플리케이션 ID를 지정하여 클러스터에 작업을 제출할 수 있습니다. 애플리케이션 ID를 가져오려면 `yarn application -list`에서 AWS CLI 또는 [YarnClient](https://hadoop.apache.org/docs/current/api/org/apache/hadoop/yarn/client/api/YarnClient.html) API 작업을 통해를 실행합니다.

```
$ yarn application -list
16/09/07 19:32:13 INFO client.RMProxy: Connecting to ResourceManager at
 ip-10-181-83-19.ec2.internal/10.181.83.19:8032
Total number of applications (application-types: [] and states: [SUBMITTED, ACCEPTED,
 RUNNING]):1
Application-Id      Application-Name      Application-Type      User      Queue      State
 Final-State      Progress      Tracking-URL
application_1473169569237_0002      Flink session with 14 TaskManagers (detached)
   Apache Flink      hadoop      default              RUNNING            UNDEFINED
 100% http://ip-10-136-154-194.ec2.internal:33089
```

이 Flink 세션의 애플리케이션 ID는 AWS CLI 또는 SDK에서 애플리케이션에 작업을 제출하는 데 사용할 수 `application_1473169569237_0002`있는 입니다.

Example SDK for Java

```
List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
    .withJar("command-runner.jar")
    .withArgs("flink", "run", "-m", "yarn-cluster", "-yid",
 "application_1473169569237_0002", "-yn", "2", "/usr/lib/flink/examples/streaming/
WordCount.jar",
      "--input", "s3://amzn-s3-demo-bucket/pg11.txt", "--output", "s3://amzn-s3-demo-
bucket/alice2/");

StepConfig flinkRunWordCount = new StepConfig()
  .withName("Flink add a wordcount step")
  .withActionOnFailure("CONTINUE")
  .withHadoopJarStep(flinkWordCountConf);
```

```
stepConfigs.add(flinkRunWordCount);

AddJobFlowStepsResult res = emr.addJobFlowSteps(new AddJobFlowStepsRequest()
   .withJobFlowId("myClusterId")
   .withSteps(stepConfigs));
```

Example AWS CLI

```
aws emr add-steps --cluster-id <j-XXXXXXXX> \
--steps Type=CUSTOM_JAR,Name=Flink_Submit_To_Long_Running,Jar=command-runner.jar,\
Args="flink","run","-m","yarn-cluster","-yid","application_1473169569237_0002",\
"/usr/lib/flink/examples/streaming/WordCount.jar",\
"--input","s3://amzn-s3-demo-bucket/pg11.txt","--output","s3://amzn-s3-demo-bucket/
alice2/" \
--region <region-code>
```

임시 Flink 작업 제출

다음 예제에서는 Flink 작업을 실행한 다음 완료 시 종료하는 임시 클러스터를 시작합니다.

Example SDK for Java

```
import java.util.ArrayList;
import java.util.List;
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;

public class Main_test {

 public static void main(String[] args) {
  AWSCredentials credentials_profile = null;
  try {
   credentials_profile = new ProfileCredentialsProvider("default").getCredentials();
  } catch (Exception e) {
   throw new AmazonClientException(
     "Cannot load credentials from .aws/credentials file. " +
       "Make sure that the credentials file exists and the profile name is specified
 within it.",
```

```
      e);
    }

    AmazonElasticMapReduce emr = AmazonElasticMapReduceClientBuilder.standard()
      .withCredentials(new AWSStaticCredentialsProvider(credentials_profile))
      .withRegion(Regions.US_WEST_1)
      .build();

    List<StepConfig> stepConfigs = new ArrayList<StepConfig>();
    HadoopJarStepConfig flinkWordCountConf = new HadoopJarStepConfig()
      .withJar("command-runner.jar")
      .withArgs("bash", "-c", "flink", "run", "-m", "yarn-cluster", "-yn", "2",
        "/usr/lib/flink/examples/streaming/WordCount.jar", "--input", "s3://path/to/
input-file.txt", "--output",
        "s3://path/to/output/");

    StepConfig flinkRunWordCountStep = new StepConfig()
      .withName("Flink add a wordcount step and terminate")
      .withActionOnFailure("CONTINUE")
      .withHadoopJarStep(flinkWordCountConf);

    stepConfigs.add(flinkRunWordCountStep);

    Application flink = new Application().withName("Flink");

    RunJobFlowRequest request = new RunJobFlowRequest()
      .withName("flink-transient")
      .withReleaseLabel("emr-5.20.0")
      .withApplications(flink)
      .withServiceRole("EMR_DefaultRole")
      .withJobFlowRole("EMR_EC2_DefaultRole")
      .withLogUri("s3://path/to/my/logfiles")
      .withInstances(new JobFlowInstancesConfig()
        .withEc2KeyName("myEc2Key")
        .withEc2SubnetId("subnet-12ab3c45")
        .withInstanceCount(3)
        .withKeepJobFlowAliveWhenNoSteps(false)
        .withMasterInstanceType("m4.large")
        .withSlaveInstanceType("m4.large"))
      .withSteps(stepConfigs);

    RunJobFlowResult result = emr.runJobFlow(request);
    System.out.println("The cluster ID is " + result.toString());
```

```
 }

}
```

Example AWS CLI

`create-cluster` 하위 명령을 사용하여 Flink 작업 완료 시 종료되는 임시 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-5.2.1 \
--name "Flink_Transient" \
--applications Name=Flink \
--configurations file://./configurations.json \
--region us-east-1 \
--log-uri s3://myLogUri \
--auto-terminate
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=<YourKeyName>,InstanceProfile=EMR_EC2_DefaultRole \
--steps Type=CUSTOM_JAR,Jar=command-runner.jar,Name=Flink_Long_Running_Session,\
Args="bash","-c","\"flink run -m yarn-cluster /usr/lib/flink/examples/streaming/
WordCount.jar
--input s3://amzn-s3-demo-bucket/pg11.txt --output s3://amzn-s3-demo-bucket/alice/""
```

Scala 셸 사용

EMR 클러스터용 Flink Scala 쉘은 새로운 YARN 세션을 시작하도록 구성됩니다. 아래 절차에 따라 Scala 셸을 사용할 수 있습니다.

프라이머리 노드에서 Flink Scala 쉘 사용

1. SSH를 사용하여 프라이머리 노드에 연결에서 설명한 대로 SSH를 사용하여 프라이머리 노드에 로그인합니다.
2. 다음을 입력하여 셸을 시작합니다.

   Amazon EMR 버전 5.5.0 이상에서는 다음 명령을 사용하여 하나의 TaskManager로 Scala 쉘용 Yarn 클러스터를 시작할 수 있습니다.

   ```
   % flink-scala-shell yarn 1
   ```

   이전 버전의 Amazon EMR에서는 다음을 사용합니다.

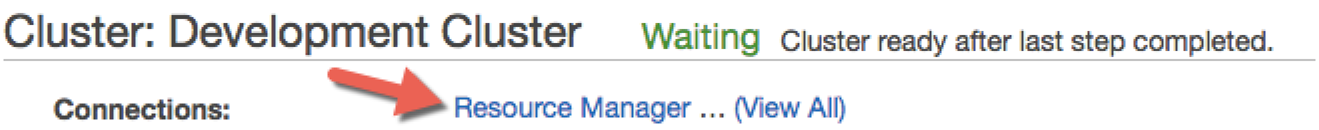
```
% /usr/lib/flink/bin/start-scala-shell.sh yarn 1
```

이렇게 하면 Flink Scala 셸이 시작되어 대화식으로 Flink를 사용할 수 있습니다. 다른 인터페이스 및 옵션과 마찬가지로, 셸에서 실행할 작업 수에 따라 이 예제에 사용된 -n 옵션 값을 조정할 수 있습니다.

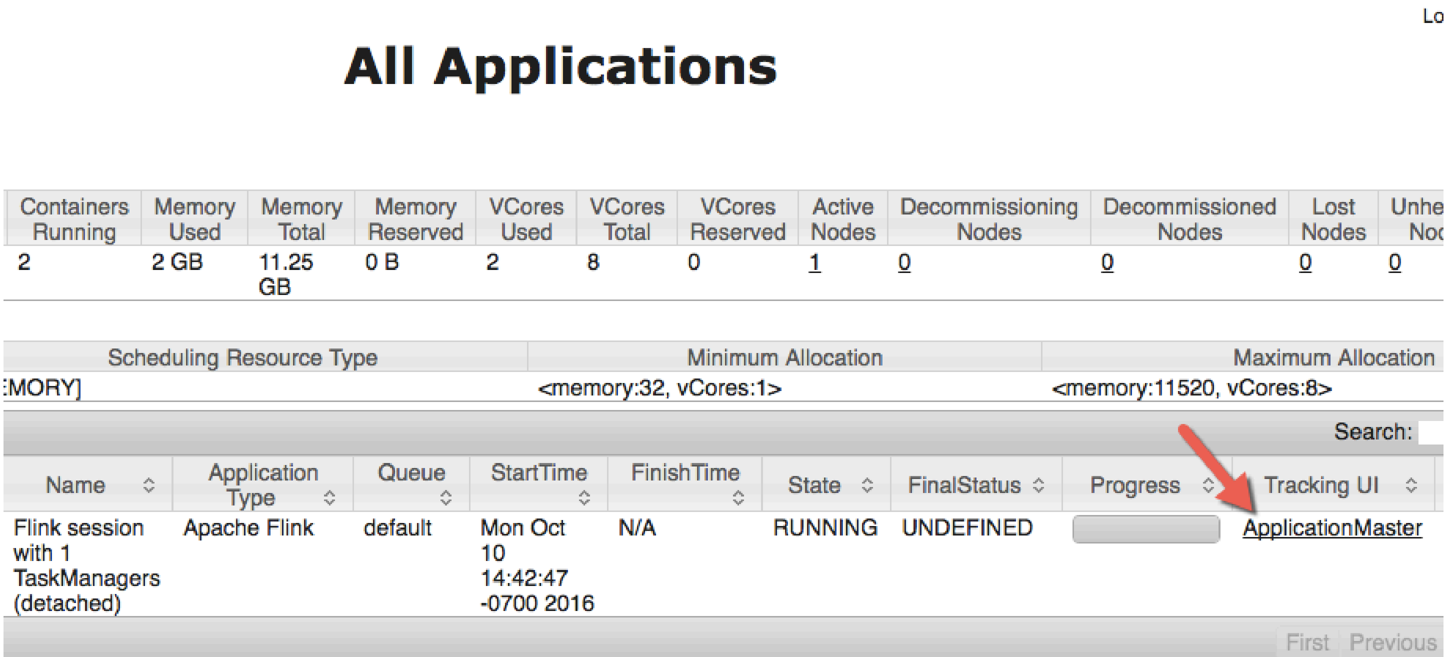
자세한 내용은 공식 Apache Flink 설명서에서 Scala REPL을 참조하세요.

Flink 웹 인터페이스 찾기

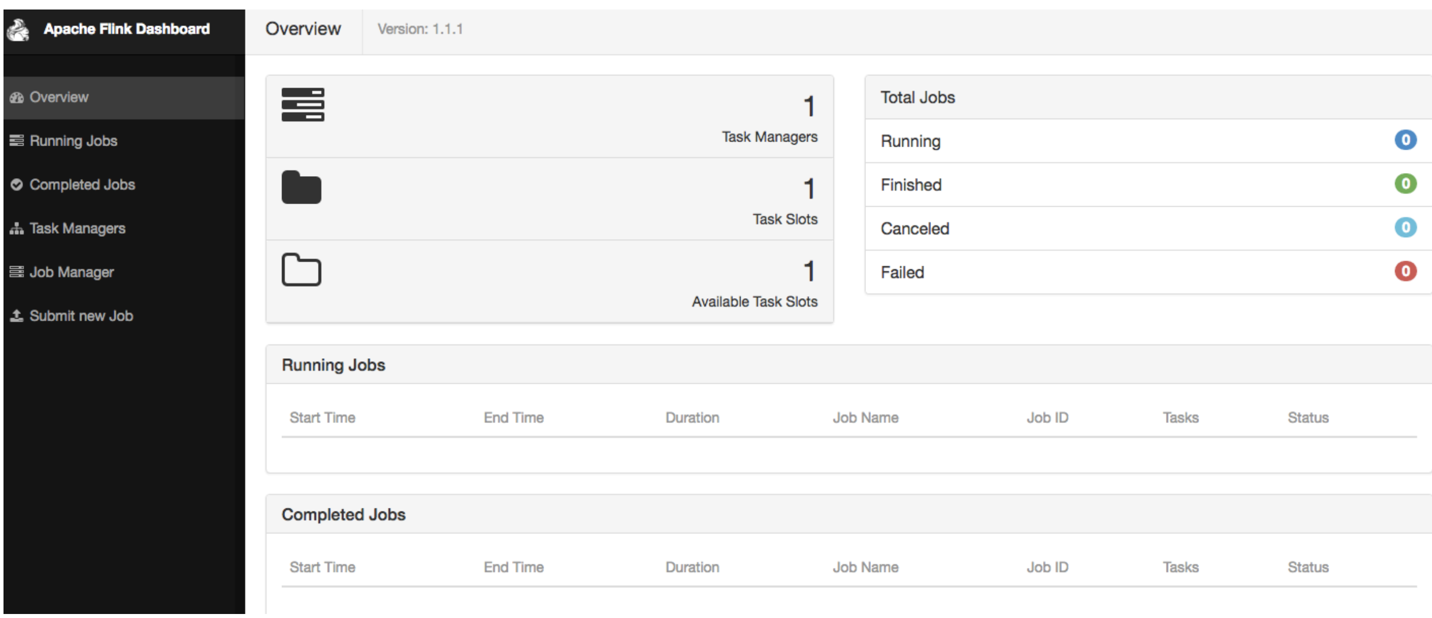
Flink 애플리케이션에 속하는 애플리케이션 마스터는 Flink 웹 인터페이스를 호스팅합니다. JAR을 작업으로 제출하거나 다른 작업의 현재 상태를 볼 수도 있습니다. Flink 웹 인터페이스는 실행 중인 Flink 세션이 있는 동안만 활성화됩니다. 장기 실행 YARN 작업이 이미 활성화되어 있는 경우, Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결 주제의 지침에 따라 YARN ResourceManager에 연결할 수 있습니다. 예를 들어 SSH 터널을 설정했으며 브라우저에서 프록시를 활성화한 경우 EMR 클러스터 세부 정보 페이지의 연결에서 ResourceManager 연결을 선택합니다.



ResourceManager를 찾은 후에는 Flink 세션을 호스팅하고 있는 YARN 애플리케이션을 선택합니다. Tracking UI(추적 UI) 열에서 링크를 선택합니다.



Flink 웹 인터페이스에서는 구성을 보고 고유의 사용자 지정 JAR를 작업으로 제출하거나 진행 중인 작업을 모니터링할 수 있습니다.



Flink 오토스케일러

개요

Amazon EMR 릴리스 6.15.0 이상에서는 Flink 오토스케일러를 지원합니다. 작업 오토스케일러 조정 기능은 실행 중인 Flink 스트리밍 작업에서 지표를 수집하고 개별 작업 버텍스의 규모를 자동 조정합니다. 그 결과 역압이 줄고 설정한 사용률 목표를 달성할 수 있습니다.

자세한 내용을 확인하려면 Apache Flink Kubernetes 연산자 문서의 [오토스케일러](#) 섹션을 참조하세요.

고려 사항

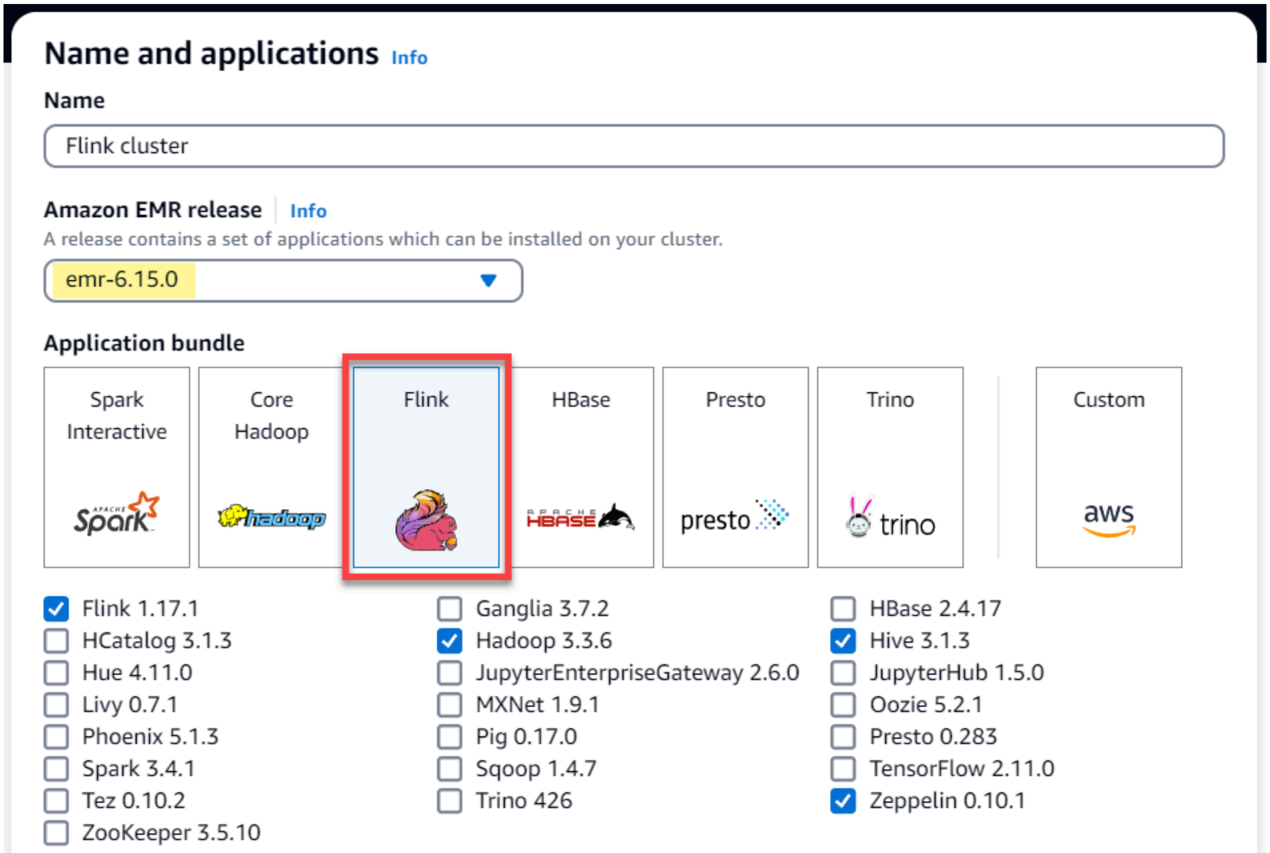
- Flink 오토스케일러는 Amazon EMR 6.15.0 이상에서 지원됩니다.
- Flink 오토스케일러는 스트리밍 작업에 대해서만 지원됩니다.
- 적응형 스케줄러만 지원됩니다. 기본 스케줄러는 지원되지 않습니다.
- 동적 리소스 프로비저닝을 위해 클러스터 스케일링 기능을 활성화할 것을 권장합니다. Amazon EMR Managed Scaling이 선호되는 이유는 5~10초마다 지표 평가가 발생하기 때문입니다. 이 간격에서 클러스터는 필요한 클러스터 리소스의 변화에 더 쉽게 적응할 수 있습니다.

오토스케일러 활성화

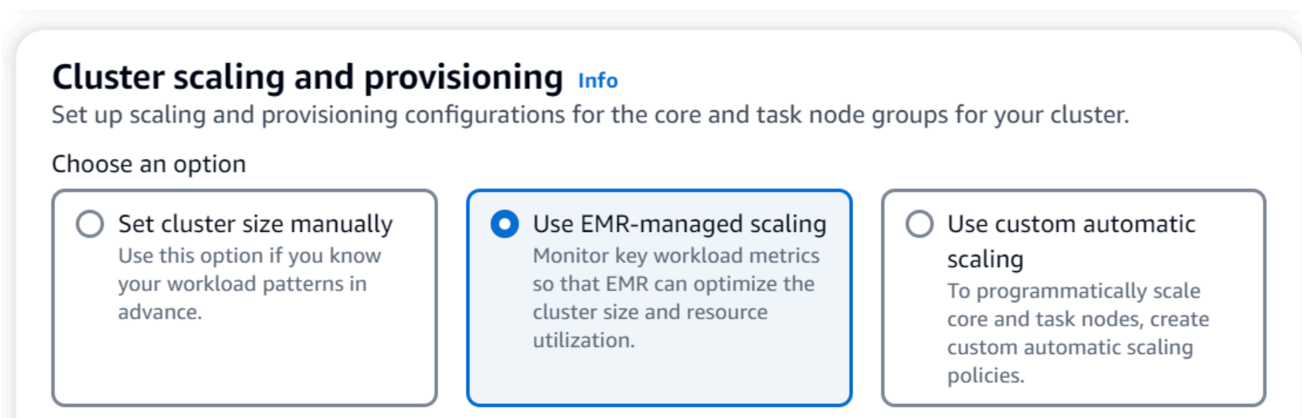
EC2 클러스터에서 Amazon EMR을 만드는 경우 다음 단계를 사용하여 Flink 오토스케일러를 활성화하세요.

1. Amazon EMR 콘솔에서 새 EMR 클러스터를 생성합니다.

    a. Amazon EMR 릴리스 `emr-6.15.0` 이상을 선택합니다. Flink 애플리케이션 번들을 선택하고 클러스터에 포함하려는 다른 애플리케이션을 선택합니다.



b. 클러스터 크기 조정 및 프로비저닝 옵션에서 EMR 관리형 크기 조정 사용을 선택합니다.



2. 소프트웨어 설정 섹션에서 다음 구성을 입력하고 Flink 오토스케일러를 활성화합니다. 테스트 시나리오의 경우에는, 결정 간격, 지표 창 간격 및 안정화 간격을 더 낮은 값으로 설정하여 작업에서 조정 결정이 즉시 내려져 검증이 수월해지도록 합니다.

```
[
  {
    "Classification": "flink-conf",
    "Properties": {
```

```
        "job.autoscaler.enabled": "true",
        "jobmanager.scheduler": "adaptive",
        "job.autoscaler.stabilization.interval": "60s",
        "job.autoscaler.metrics.window": "60s",
        "job.autoscaler.decision.interval": "10s",
        "job.autoscaler.debug.logs.interval": "60s"
      }
    }
]
```

3. 원하는 기타 설정을 선택하거나 구성한 후 Flink 오토스케일러를 사용하는 클러스터를 만드세요.

오토스케일러 구성

이 섹션에서는 특정한 요구 사항을 토대로 변경할 수 있는 구성들을 대부분 설명합니다.

> **Note**
>
> `time`, `interval` 및 `window` 설정과 같은 시간 기반 구성을 사용하는 경우, 단위가 지정되지 않은 경우의 기본 단위는 밀리초입니다. 따라서 접미사가 없는 `30` 값은 30밀리초입니다. 기타 시간 단위의 경우에는, 초의 경우 `s`, 분의 경우 `m`, 또는 시간의 경우 `h`를 적절한 접미사로 포함하세요.

주제

오토스케일러 루프 구성

오토스케일러는 몇 개의 구성 가능 시간 간격마다 작업 버텍스 레벨 지표를 가져와 실행 가능한 규모로 배율로 변환하고, 신규 작업 버텍스 병렬 처리를 예측하며, 이를 작업 스케줄러에 권장합니다. 지표는 작업 재시작 시간과 클러스터 안정화 간격 이후에만 수집됩니다.

Config 키	기본값	설명	예제 값
job.autoscaler.enabled	false	Flink 클러스터에서 오토스케일링을 활성화합니다.	true, false
job.autoscaler.decision.interval	60s	오토스케일러 결정 간격입니다.	30(기본 단위는 밀리초), 5m, 1h
job.autoscaler.restart.time	3m	운영자가 기록을 통해 확실하게 판단할 수 있을 때까지 사용하는 예상 재시작 시간입니다.	30(기본 단위는 밀리초), 5m, 1h
job.autoscaler.stabilization.interval	300s	새로운 조정이 실행되지 않는 안정화 기간입니다.	30(기본 단위는 밀리초), 5m, 1h
job.autoscaler.debug.logs.interval	300s	오토스케일러 디버그 로그 간격입니다.	30(기본 단위는 밀리초), 5m, 1h

지표 집계 및 기록 구성

오토스케일러는 지표를 가져와 시간 기반 슬라이딩 창에 대해 집계하고, 이 항목에 대한 평가를 통해 규모 조정 결정이 내려집니다. 각 작업 버텍스에 대한 규모 조정 결정 기록은 새로운 병렬 처리를 추정하는 데 활용됩니다. 이러한 항목에는 연령 기반 만료일과 기록 크기(최소 1)가 모두 포함되어 있습니다.

Config 키	기본값	설명	예제 값
job.autoscaler.metrics.window	600s	Scaling metrics aggregation window size.	30(기본 단위는 밀리초), 5m, 1h
job.autoscaler.history.max.count	3	버텍스당 유지할 과거 규모 조정 결정의 최대 수입니다.	1~Integer.MAX_VALUE

Config 키	기본값	설명	예제 값
`job.autoscaler.history.max.age`	24h	버텍스당 유지할 과거 규모 조정 결정의 최소 수입니다.	30(기본 단위는 밀리초), 5m, 1h

작업 버텍스 수준 구성

각 작업 버텍스의 병렬 처리는 목표 사용률을 근거로 수정되며 최소-최대 병렬 처리 한도에 의해 제한됩니다. 목표 사용률을 100%(즉, 값 1)에 가깝게 설정하는 것은 바람직하지 않으며 사용률 경계는 중간 부하 변동을 처리하기 위한 완충제 역할을 합니다.

Config 키	기본값	설명	예제 값
`job.autoscaler.target.utilization`	0.7	목표 버텍스 사용률.	0 - 1
`job.autoscaler.target.utilization.boundary`	0.4	목표 버텍스 사용률 경계. 현재 처리 속도가 범위 `[target_rate / (target_utilization - boundary) ~(target_rate / (target_utilization + boundary)]` 이내인 경우에는 규모 조정이 수행되지 않습니다.	0 - 1
`job.autoscaler.vertex.min-parallelism`	1	오토스케일러에서 사용할 수 있는 최소 병렬 처리입니다.	0 - 200
`job.autoscaler.vertex.max-parallelism`	200	오토스케일러에서 사용할 수 있는 최대 병렬 처리입니다. 이 한도가 Flink 구성 또는 각 운영자에 구성된 최대 병렬 처리보다 더 높은 경우 이 한도는 무시됩니다.	0 - 200

백로그 처리 구성

작업 버텍스에는 규모 조정 작업 기간 동안 누적되는 진행 중 이벤트 또는 백로그를 처리하기 위해 추가적인 리소스가 필요합니다. 이를 `catch-up` 기간이라고도 합니다. 백로그를 처리하는 데 걸리는 시간이 구성된 `lag -threshold` 값을 초과하면 작업 버텍스 목표 사용률이 최대 수준으로 높아집니다. 그 결과 백로그가 처리되는 동안 불필요한 규모 조정 작업을 막을 수 있습니다.

Config 키	기본값	설명	예제 값
`job.autoscaler.backlog-processing.lag-threshold`	5m	지연 임계값은 불필요한 크기 조정 작업을 방지하는 동시에 지연을 발생시키는 진행 중 메시지를 제거합니다.	`30`(기본 단위는 밀리초), 5m, 1h
`job.autoscaler.catch-up.duration`	15m	규모 조정 작업 후에 백로그를 완전히 처리하는 데 걸리는 목표 기간입니다. 0으로 설정하여 백로그 기반 크기 조정을 비활성화합니다.	`30`(기본 단위는 밀리초), 5m, 1h

규모 조정 작업 구성

오토스케일러는 유예 기간 내에 스케일 업 작업 직후에 스케일 다운 작업을 수행하지 않습니다. 이 경우 일시적인 부하 변동으로 인해 스케일 업-다운-업-다운 작업이 불필요하게 반복되는 것을 막을 수 있습니다.

스케일 다운 작업 비율을 사용하여 병렬 처리를 단계적으로 줄이고 리소스를 릴리스하여 일시적인 부하 급상승에 대비할 수 있습니다. 이 작업을 통해 대규모 스케일 다운 작업 후에 불필요한 소규모 스케일 업 작업이 진행되는 것을 방지할 수도 있습니다.

과거 작업 버텍스 규모 조정 결정 기록을 근거로 무의미한 스케일 업 작업을 감지하여 추가적인 변경 처리의 변경을 방지할 할 수 있습니다.

Config 키	기본값	설명	예제 값
`job.autoscaler.scale-up.grace-period`	1h	버텍스가 스케일 업된 후에 스케일 다운이 허용되지 않는 기간입니다.	`30`(기본 단위는 밀리초), 5m, 1h

Config 키	기본값	설명	예제 값
job.autoscaler.scale-down.max-factor	0.6	최대 스케일 다운 계수입니다. 값이 1이면 스케일 다운에 제한이 없음을 의미합니다. 0.6은 당초 병렬 처리의 60%로만 작업을 스케일 다운할 수 있음을 의미합니다.	0 - 1
job.autoscaler.scale-up.max-factor	100000.	최대 스케일 업 비율입니다. 값이 2.0이면 작업은 현재 병렬 처리의 200%까지만 스케일 업될 수 있습니다.	0 - Integer.MAX_VALUE
job.autoscaler.scaling.effectiveness.detection.enabled	false	무의미한 규모 조정 작업을 감지하고 자동 스케일러가 추가 스케일 업을 차단하도록 허용할지 여부입니다.	true, false

태스크 복구 및 작업 규모 조정을 위한 작업 재시작 시간 최적화

태스크가 실패하거나 규모 조정 작업이 발생할 경우 Flink는 마지막으로 완료된 체크포인트의 태스크를 다시 재실행하려고 시도합니다. 체크포인트 상태의 크기와 병렬 태스크의 수에 따라 재시작 프로세스를 실행하는 데 1분 이상 소요될 수 있습니다. 프로세스를 다시 시작하는 동안에는 작업에 대한 백로그 태스크가 누적될 수 있습니다. 그렇지만 Flink는 실행 그래프의 복구 및 재시작 속도를 최적화하여 작업 안정성을 향상시킬 수 있는 방법이 몇 가지 있습니다.

이 페이지에서는 Amazon EMR Flink를 사용하여 태스크 복구 및 규모 조정 작업에서 작업 재시작 시간을 단축할 수 있는 몇 가지 방법을 설명합니다.

주제

태스크-로컬 복구

> **Note**
>
> 태스크-로컬 복구는 Amazon EMR 6.0.0 이상에서 지원됩니다.

Flink 체크포인트를 사용할 경우 각 태스크에서 Flink가 Amazon S3와 같은 분산 스토리지에 기록하는 상태 스냅샷을 만듭니다. 복구의 경우 태스크는 분산 스토리지를 통해 해당 상태를 복원합니다. 분산 스토리지에서는 내결함성을 제공하며 모든 노드에서 액세스가 가능하기 때문에 크기 재조정이 이뤄지는 동안 상태를 재분배할 수 있습니다.

하지만 원격 분산 저장소에는 모든 태스크에서 네트워크를 통해 원격 위치에서 해당 상태를 읽어야 한다는 단점도 있습니다. 이러한 한계로 인해 태스크 복구 또는 규모 조정 작업 중에 대규모 상태의 복구 시간이 길어질 수 있습니다.

이와 같은 긴 복구 시간 문제는 태스크-로컬 복구를 통해 해결됩니다. 태스크에서는 체크포인트의 상태를 로컬 디스크와 같이 해당 작업에 대해 로컬인 보조 스토리지에 기록합니다. 또한 태스크는 기본 스토리지(이 경우 Amazon S3)에 상태를 저장합니다. 복구가 진행되는 동안 스케줄러는 태스크가 이전에 실행된 동일한 태스크 관리자에서 태스크를 예약하기 때문에 원격 상태 저장소에서 데이터를 읽는 대신 로컬 상태 저장소에서 복구할 수 있습니다. 자세한 내용을 알아보려면 Apache Flink 설명서의 [태스크 로컬 복구](https://nightlies.apache.org/flink/flink-docs-master/docs/ops/state/large_state_tuning/#task-local-recovery)를 참조하세요.

샘플 작업을 사용한 벤치마크 테스트 결과에 따르면 태스크-로컬 복구가 활성화된 상태에서는 복구 시간이 몇 분에서 몇 초로 단축된 것으로 확인되었습니다.

태스크-로컬 복구를 활성화하려면 `flink-conf.yaml` 파일에 다음과 같은 구성을 설정하세요. 체크포인트 간격의 값을 밀리초 단위로 지정하세요.

```
state.backend.local-recovery: true
state.backend: hasmap or rocksdb
state.checkpoints.dir: s3://storage-location-bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

일반 로그 기반 증분 체크포인트

> **Note**
>
> 일반 로그 기반 증분 체크포인트는 Amazon EMR 6.10.0 이상에서 지원됩니다.

체크포인트의 속도를 높이기 위해 일반 로그 기반 증분 체크포인트가 Flink 1.16에 추가되었습니다. 체크포인트 간격을 빠르게 하면 복구 후 다시 처리해야 하는 이벤트가 줄어들기 때문에 복구 작업이 적어지는 경우가 많습니다. 자세한 내용은 Apache Flink 블로그에서 [일반 로그 기반 증분 체크포인트로 체크포인트의 속도 및 안정성 강화]를 참조하세요.

샘플 작업을 이용하여 수행한 벤치마크 테스트에서 일반 로그 기반 증분 체크포인트를 사용하면 체크포인트 시간이 몇 분에서 몇 초로 단축된 것이 확인되었습니다.

일반 로그 기반 증분 체크포인트를 활성화하려면 `flink-conf.yaml` 파일에 다음 구성을 설정하세요. 체크포인트 간격의 값을 밀리초 단위로 지정하세요.

```
state.backend.changelog.enabled: true
state.backend.changelog.storage: filesystem
dstl.dfs.base-path: s3://bucket-path/changelog
state.backend.local-recovery: true
state.backend: rocksdb
state.checkpoints.dir: s3://bucket-path/checkpoint
execution.checkpointing.interval: 15000
```

세분화된 복구

> **Note**
> 기본 스케줄러에 대한 세분화된 복구 지원은 Amazon EMR 6.0.0에서 이용할 수 있습니다. 적응형 스케줄러에서의 세분화된 복구 지원은 Amazon EMR 6.15.0 이상에서 이용할 수 있습니다.

실행 중에 태스크가 실패하면 Flink에서는 전체 실행 그래프를 재설정하고 마지막으로 완료된 체크포인트에서 전체 재실행을 트리거합니다. 이 방식은 실패한 태스크를 재실행하는 것보다 비용이 많이 듭니다. 세분화된 복구에서는 실패한 태스크의 파이프라인으로 연결된 구성 요소만 재시작합니다. 다음 예제의 작업 그래프에는 버텍스가 5개(A~E) 있습니다. 버텍스 사이에 있는 모든 연결은 포인트별 분포로 파이프라인되며 작업의 `parallelism.default`는 2로 설정됩니다.

```
A # B # C # D # E
```

이 예시에서는 태스크가 총 10개 실행 중입니다. 첫 번째 파이프라인(a1~e1)은 TaskManager(TM1)에서 실행되고 두 번째 파이프라인(a2~e2)은 또 다른 TaskManager(TM2)에서 실행됩니다.

```
a1 # b1 # c1 # d1 # e1
a2 # b2 # c2 # d2 # e2
```

`a1 # e1` 및 `a2 # e2`라는 두 개의 구성 요소가 파이프라인으로 연결되어 있습니다. `TM1` 또는 `TM2` 중에 하나가 실패하면 실패는 TaskManager가 중이던 파이프라인에 있는 태스크 5개에만 영향을 미칩니다. 재시작 전략에 따라 영향을 받는 파이프라인 구성 요소만 시작됩니다.

세분화된 복구는 완벽히 병렬화된 Flink 작업에서만 작동합니다. `keyBy()` 또는 `redistribute()` 작업에서는 지원되지 않습니다. 자세한 내용은 Flink 개선 제안 Jira 프로젝트의 FLIP-1: 태스크 실패에서 세분화된 복구를 참조하세요.

세분화된 복구를 활성화하려면 `flink-conf.yaml` 파일에 다음과 같은 구성을 설정하세요.

```
jobmanager.execution.failover-strategy: region
restart-strategy: exponential-delay or fixed-delay
```

적응형 스케줄러의 결합된 재시작 메커니즘

> **Note**
>
> 적응형 스케줄러의 결합된 재시작 메커니즘은 Amazon EMR 6.15.0 이상에서 지원됩니다.

적응형 스케줄러에서는 가용 슬롯을 기반으로 작업 병렬성을 조정할 수 있습니다. 이 스케줄러는 구성된 작업 병렬 처리에 적합한 가용 슬롯이 충분하지 않은 경우 병렬 처리의 수를 자동으로 줄입니다. 새 슬롯이 가용 상태가 되면 작업은 구성된 작업 병렬 처리로 다시 확장됩니다. 적응형 스케줄러는 가용 리소스가 충분하지 않은 경우 작업에서 가동 중지가 발생하는 것은 방지합니다. Flink Autoscaler에 대해 지원되는 스케줄러입니다. 이러한 이유들로 인해 적응형 스케줄러를 Amazon EMR Flink와 함께 사용하는 것이 좋습니다. 단, 적응형 스케줄러는 짧은 시간 내에 여러 번 재시작을 수행할 수 있으며, 새 리소스가 추가될 때마다 한 번씩 다시 시작됩니다. 이로 인해 작업 성능이 떨어질 수 있습니다.

Amazon EMR 6.15.0 이상에서는 Flink에 첫 번째 리소스가 추가될 때 재시작 기간을 연 다음, 구성된 기본 1분 간격까지 기다리는 적응형 스케줄러의 결합된 재시작 메커니즘이 있습니다. 이 메커니즘에서는 구성된 병렬 처리로 작업을 실행하기 위한 가용 리소스가 충분하거나 간격 제한 시간이 초과될 경우 단일 재시작을 수행합니다.

샘플 작업을 이용한 벤치마크 테스트에서는 적응형 스케줄러와 Flink Autoscaler를 사용할 경우 이 기능이 기본 동작보다 10% 더 많은 레코드를 처리하는 것이 입증되었습니다.

결합된 재시작 메커니즘을 활성화하려면 `flink-conf.yaml` 파일에 다음 구성을 설정하세요.

```
jobmanager.adaptive-scheduler.combined-restart.enabled: true
jobmanager.adaptive-scheduler.combined-restart.window-interval: 1m
```

Amazon EMR의 Zeppelin에서 Flink 작업

소개

Amazon EMR 릴리스 6.10.0 이상에서는 Apache Flink와 [Apache Zeppelin](https://zeppelin.apache.org/)의 통합을 지원합니다. Zeppelin 노트북을 통해 대화형 방식으로 Flink 작업을 제출할 수 있습니다. Flink 인터프리터를 사용하면 Flink 쿼리를 실행하고, Flink 스트리밍 및 배치 작업을 정의하며, Zeppelin 노트북 내에서 출력을 시각화할 수 있습니다. Flink 인터프리터는 Flink REST API를 기반으로 빌드됩니다. 이를 통해 Zeppelin 환경 내에서 Flink 작업에 액세스하고 해당 작업을 조작하여 실시간 데이터 처리 및 분석을 수행할 수 있습니다.

Flink 인터프리터에는 4개의 하위 인터프리터가 있습니다. 용도는 다양하지만 모두 JVM에 속하며 Flink에 사전 구성된 동일한 진입점(`ExecutionEnviroment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment`)을 공유합니다. 인터프리터는 다음과 같습니다.

- `%flink` - `ExecutionEnvironment`, `StreamExecutionEnvironment`, `BatchTableEnvironment`, `StreamTableEnvironment`를 생성하고 Scala 환경 제공
- `%flink.pyflink` - Python 환경 제공
- `%flink.ssql` - 스트리밍 SQL 환경 제공
- `%flink.bsql` - 배치 SQL 환경 제공

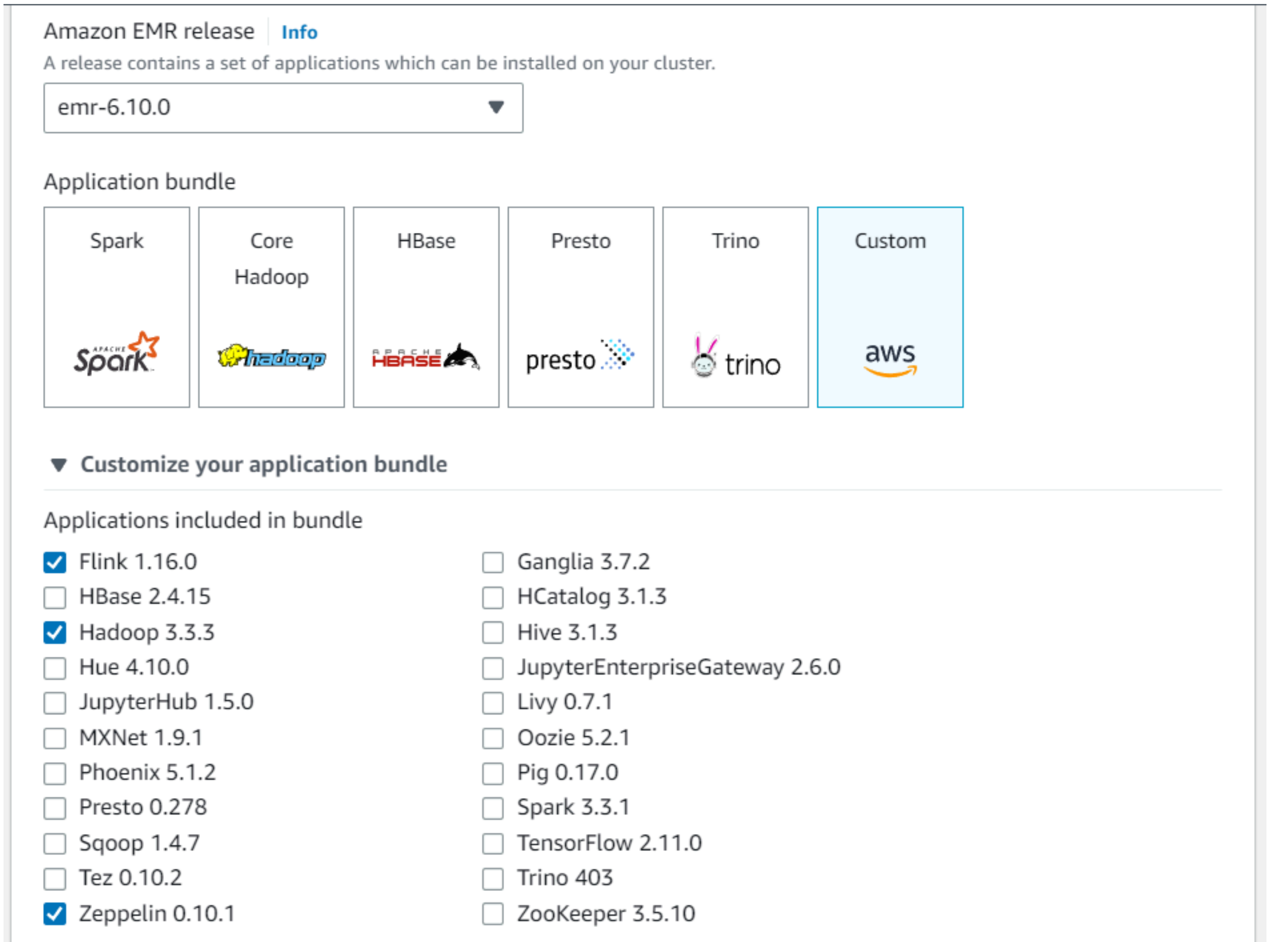
사전 조건

- Amazon EMR 6.10.0 이상에서 생성된 클러스터에 대해 Flink와 Zeppelin의 통합이 지원됩니다.
- 이러한 단계에 필요한 EMR 클러스터에 호스팅되는 웹 인터페이스를 보려면 인바운드 액세스를 허용하도록 SSH 터널을 구성해야 합니다. 자세한 내용은 [프라이머리 노드에서 호스팅되는 웹 사이트를 표시하도록 프록시 설정 구성](#)을 참조하세요.

EMR 클러스터에서 Zeppelin-Flink 구성

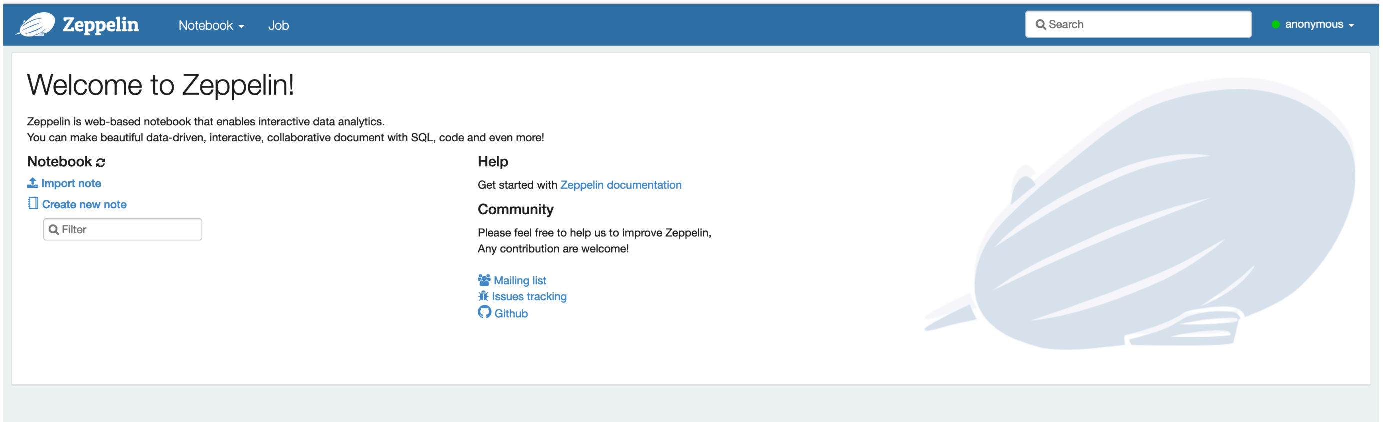
다음 단계를 사용하여 Apache Zeppelin 기반 Apache Flink가 EMR 클러스터에서 실행되도록 구성합니다.

1. Amazon EMR 콘솔에서 새 클러스터를 생성합니다. Amazon EMR 릴리스의 경우 emr-6.10.0 이상을 선택합니다. 그런 다음 사용자 지정 옵션을 사용하여 애플리케이션 번들을 사용자 지정하도록 선택합니다. 번들에 최소한 Flink, Hadoop, Zeppelin을 포함합니다.

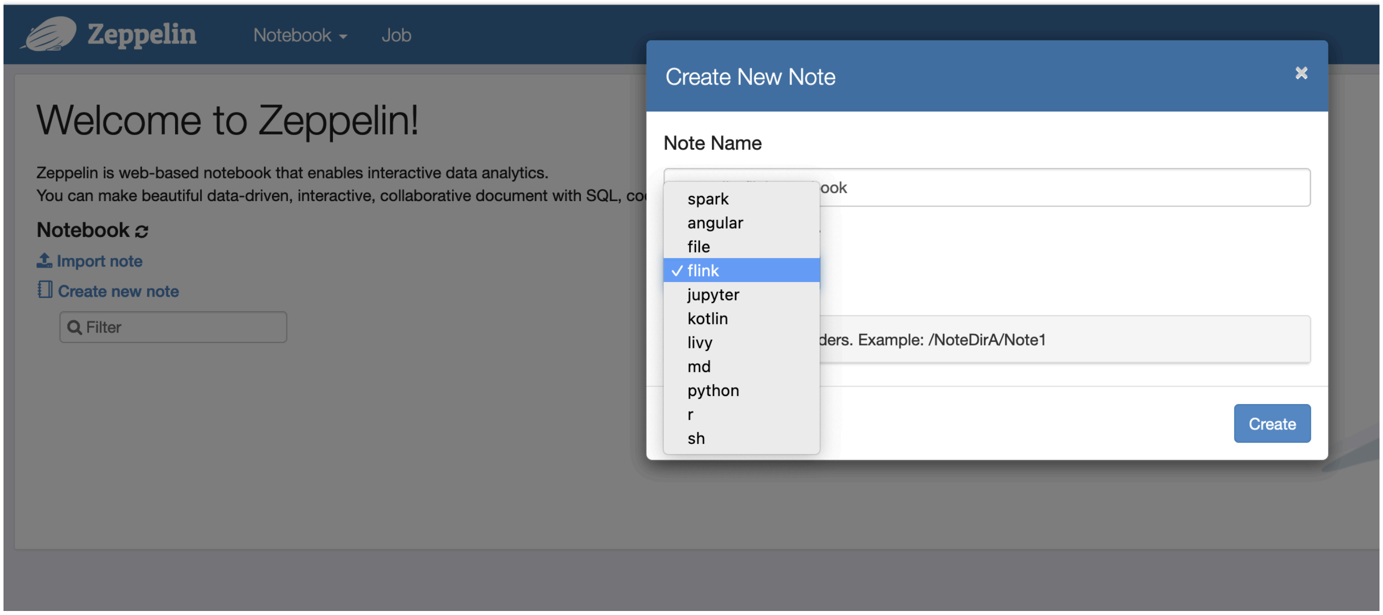


2. 원하는 설정으로 나머지 클러스터를 생성합니다.

3. 클러스터가 실행되면 콘솔에서 클러스터를 선택하여 세부 정보를 확인하고 애플리케이션 탭을 엽니다. 애플리케이션 사용자 인터페이스 섹션에서 Zeppelin을 선택하여 Zeppelin 웹 인터페이스를 엽니다. 사전 조건에 설명된 대로 프라이머리 노드에 대한 SSH 터널과 프록시 연결을 사용하여 Zeppelin 웹 인터페이스에 대한 액세스를 설정했는지 확인합니다.



4. 이제 Flink를 기본 인터프리터로 사용하여 Zeppelin 노트북에서 새 노트를 생성할 수 있습니다.



5. Zeppelin 노트북에서 Flink 작업을 실행하는 방법을 보여주는 다음 코드 예제를 참조하세요.

EMR 클러스터에서 Zeppelin-Flink를 사용하여 Flink 작업 실행

- 예제 1, Flink Scala

a) 배치 WordCount 예제(SCALA)

```
%flink

val data = benv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
             .map(w => (w, 1))
```

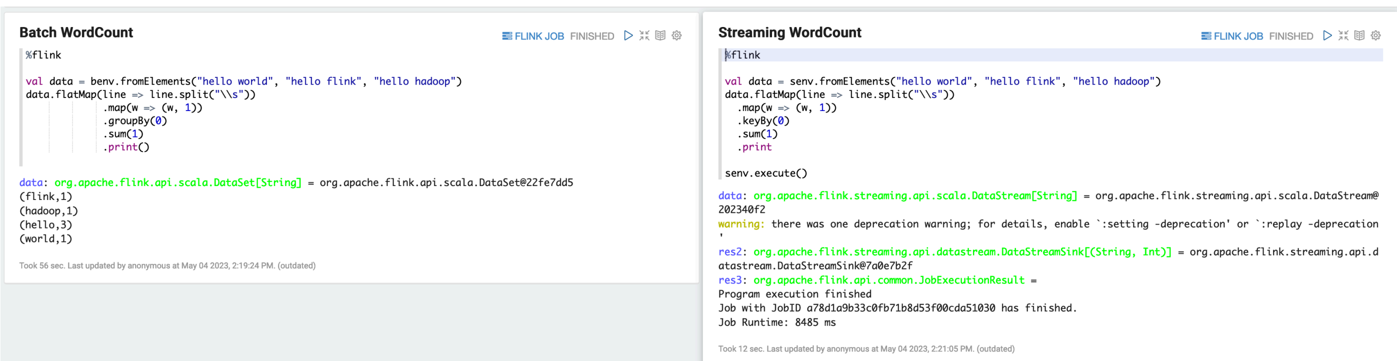
```
              .groupBy(0)
              .sum(1)
              .print()
```

b) WordCount 스트리밍 예제(SCALA)

```
%flink

val data = senv.fromElements("hello world", "hello flink", "hello hadoop")
data.flatMap(line => line.split("\\s"))
  .map(w => (w, 1))
  .keyBy(0)
  .sum(1)
  .print

senv.execute()
```



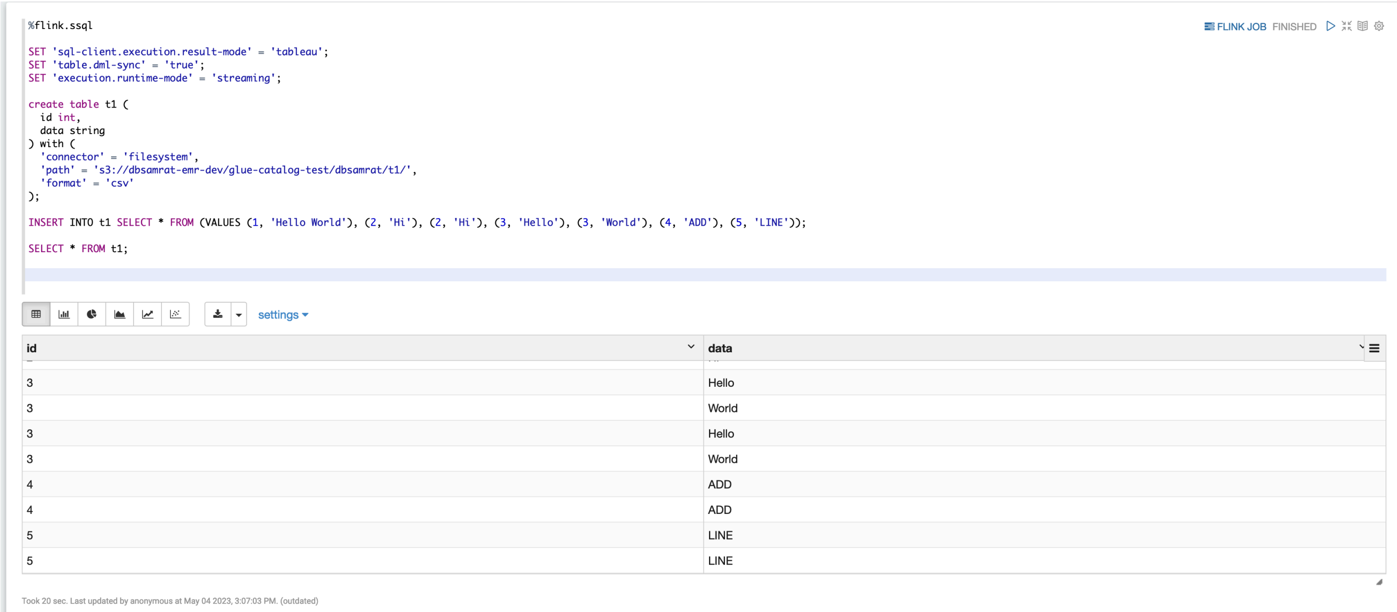
- 예제 2, Flink 스트리밍 SQL

```
%flink.ssql
SET 'sql-client.execution.result-mode' = 'tableau';
SET 'table.dml-sync' = 'true';
SET 'execution.runtime-mode' = 'streaming';

create table dummy_table (
  id int,
  data string
) with (
  'connector' = 'filesystem',
  'path' = 's3://s3-bucket/dummy_table',
  'format' = 'csv'
);
```

```
INSERT INTO dummy_table SELECT * FROM (VALUES (1, 'Hello World'), (2, 'Hi'), (2,
 'Hi'), (3, 'Hello'), (3, 'World'), (4, 'ADD'), (5, 'LINE'));

SELECT * FROM dummy_table;
```



- 예제 3, Pyflink. 샘플 텍스트 파일 `word.txt`를 S3 버킷에 업로드해야 합니다.

```
%flink.pyflink

import argparse
import logging
import sys

from pyflink.common import Row
from pyflink.table import (EnvironmentSettings, TableEnvironment, TableDescriptor,
 Schema,
                           DataTypes, FormatDescriptor)
from pyflink.table.expressions import lit, col
from pyflink.table.udf import udtf

def word_count(input_path, output_path):
    t_env = TableEnvironment.create(EnvironmentSettings.in_streaming_mode())
    # write all the data to one file
    t_env.get_config().set("parallelism.default", "1")

    # define the source
    if input_path is not None:
```

```
        t_env.create_temporary_table(
            'source',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .build())
                           .option('path', input_path)
                           .format('csv')
                           .build())
        tab = t_env.from_path('source')
    else:
        print("Executing word_count example with default input data set.")
        print("Use --input to specify file input.")
        tab = t_env.from_elements(map(lambda i: (i,), word_count_data),
                                  DataTypes.ROW([DataTypes.FIELD('line',
 DataTypes.STRING())]))

    # define the sink
    if output_path is not None:
        t_env.create_temporary_table(
            'sink',
            TableDescriptor.for_connector('filesystem')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .option('path', output_path)
                           .format(FormatDescriptor.for_format('canal-json')
                                   .build())
                           .build())
    else:
        print("Printing result to stdout. Use --output to specify output path.")
        t_env.create_temporary_table(
            'sink',
            TableDescriptor.for_connector('print')
                           .schema(Schema.new_builder()
                                   .column('word', DataTypes.STRING())
                                   .column('count', DataTypes.BIGINT())
                                   .build())
                           .build())

    @udtf(result_types=[DataTypes.STRING()])
    def split(line: Row):
        for s in line[0].split():
```

```
            yield Row(s)

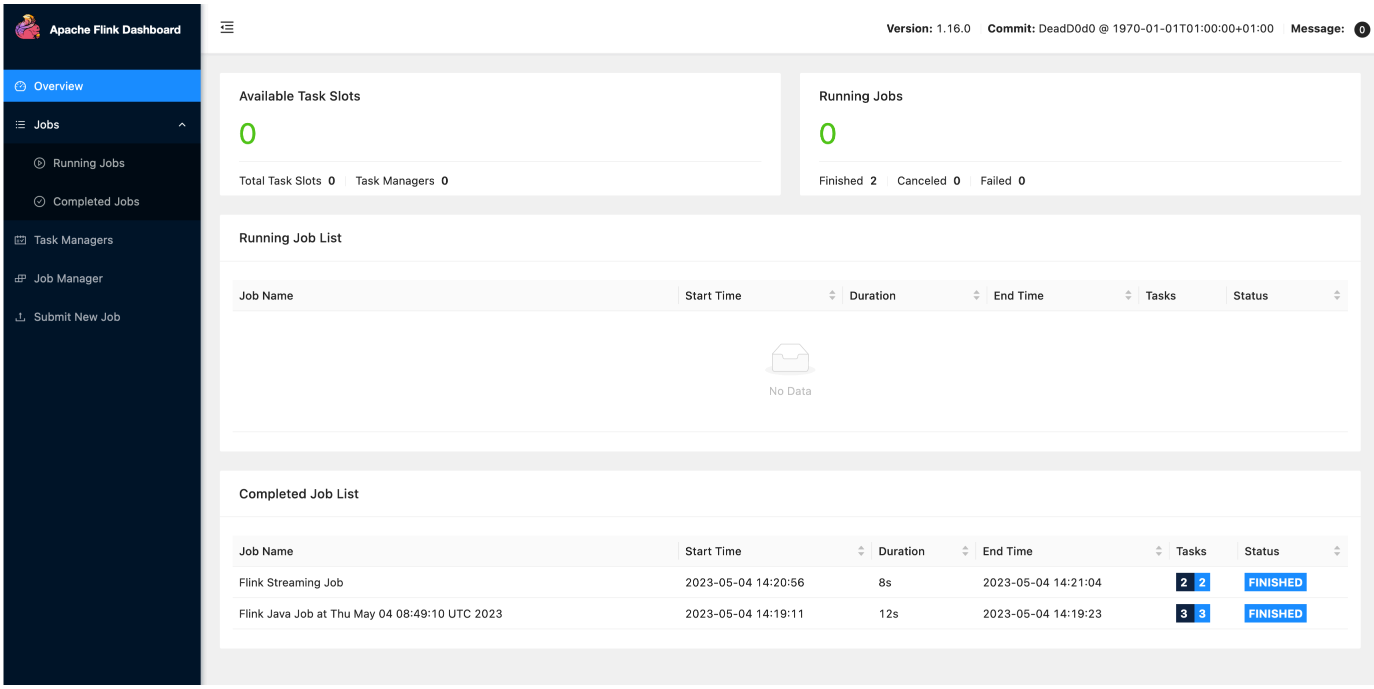
    # compute word count
    tab.flat_map(split).alias('word') \
       .group_by(col('word')) \
       .select(col('word'), lit(1).count) \
       .execute_insert('sink') \
       .wait()


logging.basicConfig(stream=sys.stdout, level=logging.INFO, format="%(message)s")


word_count("s3://s3_bucket/word.txt", "s3://s3_bucket/demo_output.txt")
```

1. Zeppelin UI에서 FLINK JOB을 선택하여 Flink 웹 UI에 액세스하고 이를 확인합니다.

2. FLINK JOB 경로를 선택하여 브라우저의 다른 탭에서 Flink 웹 콘솔로 이동합니다.



Flink 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Flink의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Flink 버전 정보

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-7.9.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client,

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
		flink-jobmanager-config, hudi, delta
emr-7.8.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.7.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-7.6.0	1.20.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.5.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.4.0	1.19.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-7.3.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.2.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-5.36.2	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-7.1.0	1.18.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta
emr-7.0.0	1.18.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.15.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-6.14.0	1.17.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.13.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.12.0	1.17.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-6.11.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.11.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi, delta-standalone-connectors
emr-6.10.1	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-6.10.0	1.16.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.9.1	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.9.0	1.15.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-6.8.1	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.8.0	1.15.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.7.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.36.1	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.36.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.6.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.35.0	1.14.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.5.0	1.14.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi
emr-6.4.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config, hudi

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-6.3.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.3.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.2.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-6.2.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-6.1.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-6.1.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.34.0	1.13.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.33.1	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.33.0	1.12.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.32.1	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.32.0	1.11.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.31.1	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.31.0	1.11.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client, flink-jobmanager-config
emr-5.30.2	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.30.1	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.30.0	1.10.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.29.0	1.9.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.28.1	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.28.0	1.9.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.27.1	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.27.0	1.8.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.26.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.25.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.24.1	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.24.0	1.8.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.23.1	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.23.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.22.0	1.7.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.2	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.21.1	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.21.0	1.7.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.20.1	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.20.0	1.6.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.19.1	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.19.0	1.6.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.18.1	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.18.0	1.6.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.2	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.17.1	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.17.0	1.5.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.16.1	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.16.0	1.5.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.15.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.15.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.2	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.14.1	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.14.0	1.4.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.13.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.13.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.3	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.2	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.12.1	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.12.0	1.4.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.4	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.11.3	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.2	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.11.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.11.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.10.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.10.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.9.1	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.9.0	1.3.2	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.3	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.8.2	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.1	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.8.0	1.3.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.7.1	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.7.0	1.3.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.6.1	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.6.0	1.2.1	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, flink-client
emr-5.5.4	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.3	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.5.2	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.5.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.4.1	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.4.0	1.2.0	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.2	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.1	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.3.0	1.1.4	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.2.3	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.2	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.2.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

Amazon EMR 릴리스 레이블	Flink 버전	Flink와 함께 설치된 구성 요소
emr-5.1.1	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client
emr-5.1.0	1.1.3	emrfs, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, flink-client

버전별 Flink 릴리스 정보

전체 릴리스 정보는 다음 섹션을 참조하세요.

Amazon EMR 7.9.0 - Flink 릴리스 정보

Amazon EMR 7.9.0 - Flink 변경 사항

유형	설명
새로운 기능	Amazon EMR 7.9.0부터 Apache Flink는 Avro, Parquet 및 ORC 파일 형식을 out-of-the-box 지원합니다. 추가 구성 없이 모든 Flink API(DataStream, Table 또는 SQL)에서 이러한 형식을 직접 사용할 수 있습니다.

유형	설명
새로운 기능	Amazon EMR 버전 7.9.0부터 구성 설정을 사용하여 Hive 메타스토어 또는 AWS Glue 데이터 카탈로그를 더 쉽게 활성화할 수 있습니다. 클러스터 생성 중에 `flink-conf` 분류`glue.enabled: true` 에 `hive.enabled: true` 또는 를 추가하여 각 데이터 카탈로그를 자동으로 구성합니다. 이렇게 간소화된 접근 방식은 이전에 필요했던 수동 구성 단계를 제거합니다.

Amazon EMR 7.8.0 - Flink 릴리스 정보

구성 - EMR Flink는 모든 AWS 리전/파티션에서 S3A와 함께 즉시 작동합니다.

Amazon EMR 7.7.0 - Flink 릴리스 정보

- Flink SQL 쉘`flink-sql-client`은에 대칭 연결된 명령을 사용하여 쉽게 호출할 수 있습니다. `/usr/lib/flink/bin/sql-client.sh`

Amazon EMR 7.6.0 - Flink 릴리스 정보

Amazon EMR 7.6.0 - Flink 기능

- 릴리스에는 변경 사항이 없습니다.

Amazon EMR 7.5.0 - Flink 릴리스 정보

유형	설명
Feature	원격 jar로 Flink 작업을 실행하기 위한 지원이 추가되었습니다.
개선	버텍스 제외 및 포함 스레드를 안전하게 만듭니다.

Amazon EMR 7.5.0 - Flink 기능

- Amazon EMR 7.5.0부터 `run` 및 `run-application` Apache Flink CLI 명령을 사용할 때 Amazon S3 위치를 JAR 경로로 지정할 수 있습니다. S3 경로를 제공하면 EMR은 Amazon S3에서 클러스터의 EBS 스토리지로 JAR 파일을 자동으로 다운로드합니다. 동일한 JAR 파일을 지정할 때마다 EMR은 클러스터의 기존 JAR 파일을 재사용하는 대신 Amazon S3에서 최신 버전을 다운로드합니다.
- Amazon EMR 7.5.0부터 고객은 `run` 및 `run-application` Flink CLI 명령을 사용하여 원격 경로(S3 위치)를 JAR 경로로 전달할 수 있습니다. 그러면 JAR이 S3 스토어에서 클러스터의 EBS 스토리지로 자동으로 가져옵니다. 동일한 JAR이 다시 제공되면 S3에서 최신 JAR을 다운로드하고 클러스터에서 기존 JAR을 재사용하지 않습니다.

Amazon EMR 7.4.0 - Flink 릴리스 정보

유형	설명
업그레이드	Flink 버전이 1.19.1로 업그레이드되었습니다.

Amazon EMR 7.3.0 - Flink 릴리스 정보

- 기본적으로 보안 구성을 통해 전송 중 데이터 암호화가 활성화된 클러스터는 Flink 프로세스, 작업 관리자 REST 엔드포인트 및 Flink 작업 기록 서버 간 내부 통신에 TLS 1.3을 사용합니다.

Amazon EMR 7.2.0 - Flink 릴리스 정보

유형	설명
개선 사항	`kubernetes.service.labels` 구성을 통해 Flink 작업당 Kubernetes 서비스에 사용자 지정 레이블 추가를 지원합니다.

Ganglia

> **Note**
>
> Ganglia를 포함하는 Amazon EMR의 마지막 릴리스는 Amazon EMR 6.15.0이었습니다. 클러스터를 모니터링하기 위해 6.15.0 이상의 릴리스에는 Amazon CloudWatch 에이전트가 포함됩니다.

Ganglia 오픈 소스 프로젝트는 성능에 미치는 영향을 최소화하면서 클러스터 및 그리드를 모니터링하도록 설계된 확장 가능한 분산 시스템입니다. 클러스터에서 Ganglia를 활성화하면 보고서를 생성하고 전체 클러스터의 성능을 볼 수 있으며 개별 노드 인스턴스의 성능을 검사할 수 있습니다. 또한 Ganglia는 Hadoop 및 Spark 지표를 수집하고 시각화하도록 구성되어 있습니다. Ganglia 오픈 소스 프로젝트에 대한 자세한 내용은 http://ganglia.info/를 참조하십시오.

브라우저에서 Ganglia 웹 UI를 보면 로드, 메모리 사용량, CPU 사용률 및 클러스터의 네트워크 트래픽을 자세히 설명하는 그래프와 함께 클러스터 성능의 개요를 볼 수 있습니다. 클러스터 통계 아래에는 클러스터의 각 개별 서버에 대한 그래프가 있습니다.

다음 테이블에는 Amazon EMR이 Ganglia를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Ganglia의 버전이 나열되어 있습니다.

이 릴리스에서 Ganglia와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Ganglia 버전 정보

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.15.0	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
		cemanager, hadoop-yarn-timeline-server, webserver

다음 테이블에는 Amazon EMR이 Ganglia를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Ganglia의 버전이 나열되어 있습니다.

이 릴리스에서 Ganglia와 함께 설치된 구성 요소의 버전은 [릴리스 5.36.2 구성 요소 버전]을 참조하세요.

emr-5.36.2용 Ganglia 버전 정보

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.36.2	Ganglia 3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

주제

- [Ganglia를 포함하는 클러스터 생성]
- [Ganglia 지표 보기]
- [Ganglia에서 Hadoop 및 Spark 지표]
- [Ganglia 릴리스 기록]

Ganglia를 포함하는 클러스터 생성

> **Note**
>
> Ganglia를 포함하는 Amazon EMR의 마지막 릴리스는 Amazon EMR 6.15.0이었습니다. 클러스터를 모니터링하기 위해 6.15.0 이상의 릴리스에는 Amazon CloudWatch 에이전트가 포함됩니다.

콘솔을 사용하여 Ganglia가 설치된 클러스터를 생성하려면

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성을 선택합니다.
3. 소프트웨어 구성에서 모든 애플리케이션 또는 Core Hadoop(코어 하둡) 또는 Spark를 선택합니다.
4. 계속해서 적절한 구성으로 클러스터를 생성합니다.

를 사용하여 클러스터에 Ganglia를 추가하려면 AWS CLI

에서 `--applications` 파라미터와 `create-cluster` 함께를 사용하여 클러스터에 Ganglia를 추가할 AWS CLI수 있습니다. `--applications`파라미터를 사용하여 Ganglia만 지정하면, Ganglia만 설치됩니다.

- 클러스터를 생성할 때 Ganglia를 추가하려면 다음 명령을 입력하고 *myKey*에 해당 EC2 키 페어의 이름을 지정합니다.

  > **Note**
  >
  > 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

  ```
  aws emr create-cluster --name "Spark cluster with Ganglia" --release-label
   emr-6.15.0; \
  --applications Name=Spark Name=Ganglia \
  --ec2-attributes KeyName=myKey --instance-type m5.xlarge \
  --instance-count 3 --use-default-roles
  ```

`--instance-groups` 파라미터를 사용하지 않고 인스턴스 수를 지정하면 단일 마스터 노드가 시작되고 나머지 인스턴스는 코어 노드로 시작됩니다. 모든 노드에는 이 명령에 지정된 인스턴스 유형이 사용됩니다.

> **Note**
>
> EMR 서비스 역할과 EC2 인스턴스 프로파일을 아직 생성하지 않았다면 `aws emr create-default-roles` 하위 명령을 입력하기 전에 `create-cluster`를 입력하여 생성합니다.

에서 Amazon EMR 명령을 사용하는 방법에 대한 자세한 내용은 섹션을 AWS CLI참조하세요https://docs.aws.amazon.com/cli/latest/reference/emr.

Ganglia 지표 보기

> **Note**
>
> Ganglia를 포함하는 Amazon EMR의 마지막 릴리스는 Amazon EMR 6.15.0이었습니다. 클러스터를 모니터링하기 위해 6.15.0 이상의 릴리스에는 Amazon CloudWatch 에이전트가 포함됩니다.

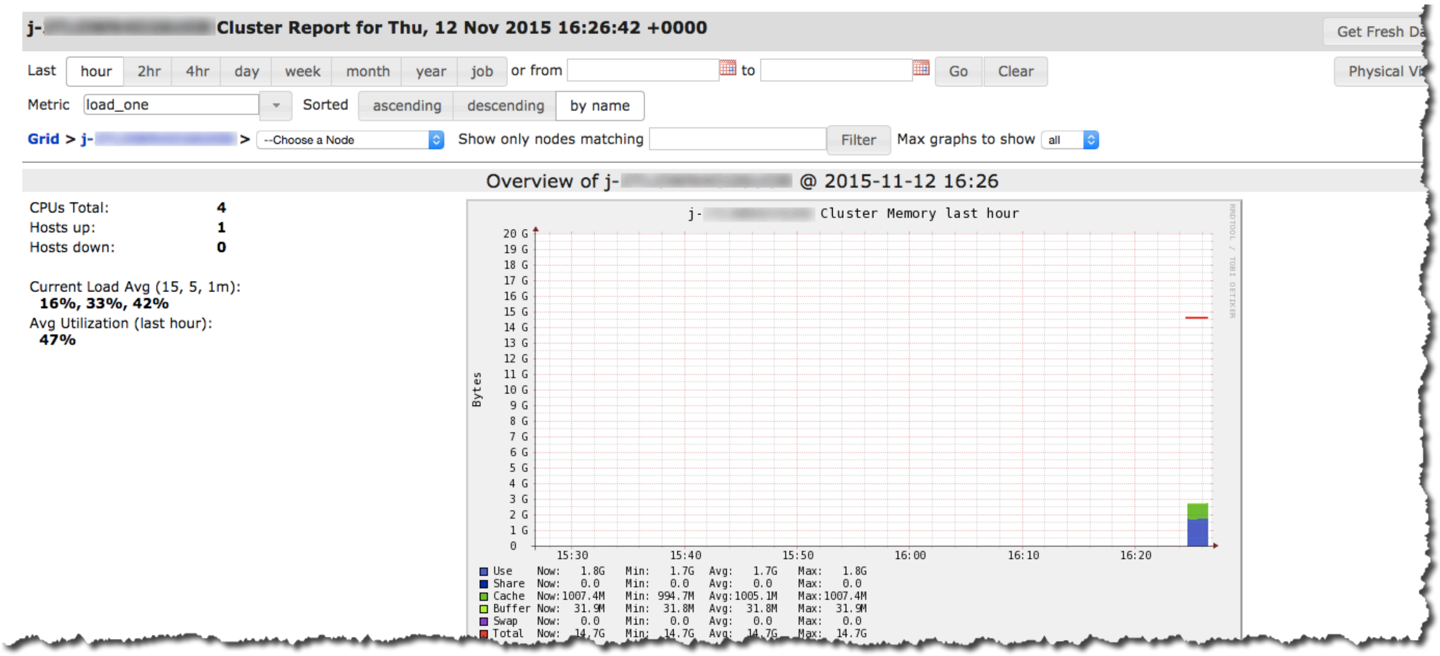
Ganglia는 Ganglia가 수집하는 지표를 볼 수 있는 웹 기반 사용자 인터페이스를 제공합니다. Amazon EMR에서 Ganglia를 실행하면 웹 인터페이스가 프라이머리 노드에서 실행되며 SSH 터널 생성이라고도 하는 포트 전달을 사용하여 볼 수 있습니다. Amazon EMR에서 웹 인터페이스를 보는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 EMR 클러스터에 호스팅된 웹 인터페이스 보기를 참조하세요.

Ganglia 웹 인터페이스를 보려면

1. SSH를 사용하여 마스터 노드로 터널링하고 보안 연결을 작성하십시오. 프라이머리 노드로의 SSH 터널을 생성하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 옵션 2, 1부: 동적 포트 전달을 사용하여 프라이머리 노드에 대해 SSH 터널 설정을 참조하세요.
2. *ec2*.amazonaws.com* 유형의 도메인에서 SOCKS 프록시를 생성하려면 Firefox용 FoxyProxy 플러그인과 같은 프록시 도구를 사용하여 웹 브라우저를 설치합니다. 자세한 내용은 Amazon

EMR 관리 안내서에서 [옵션 2, 2부: 프라이머리 노드에 호스팅된 웹 사이트를 볼 수 있도록 프록시 설정 구성]을 참조하세요.

3. 프록시 세트와 SSH 연결이 열린 상태에서 http://*master-public-dns-name*/ganglia/가 포함된 브라우저 창을 열어서 Ganglia UI를 볼 수 있습니다. 여기서 *master-public-dns-name*은 EMR 클러스터의 마스터 서버의 퍼블릭 DNS 주소입니다.



Ganglia에서 Hadoop 및 Spark 지표

> **Note**
>
> Ganglia를 포함하는 Amazon EMR의 마지막 릴리스는 Amazon EMR 6.15.0이었습니다. 클러스터를 모니터링하기 위해 6.15.0 이상의 릴리스에는 [Amazon CloudWatch 에이전트]가 포함됩니다.

Ganglia는 각 인스턴스에 대한 Hadoop 지표를 보고합니다. 분산 파일 시스템(dfs.*), Java 가상 머신(jvm.*), MapReduce(mapred.*), 원격 프로시저 호출(rpc.*) 등의 범주가 여러 유형의 지표 앞에 옵니다.

Spark 및 하둡과 같은 YARN 기반의 Ganglia 지표는 EMR 릴리스 버전 4.4.0 및 4.5.0에서 사용할 수 없습니다. 이러한 지표를 사용하려면 이후 버전을 사용하십시오.

Spark의 Ganglia 지표에는 일반적으로 YARN 애플리케이션 ID와 Spark DAGScheduler의 접두사가 있습니다. 따라서 접두어는 다음 형식을 따릅니다.

- DAGScheduler.*
- application_xxxxxxxxxx_xxxx.driver.*
- application_xxxxxxxxxx_xxxx.executor.*

Ganglia 릴리스 기록

> **Note**
>
> Ganglia를 포함하는 Amazon EMR의 마지막 릴리스는 Amazon EMR 6.15.0이었습니다. 클러스터를 모니터링하기 위해 6.15.0 이상의 릴리스에는 Amazon CloudWatch 에이전트가 포함됩니다.

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Ganglia의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Ganglia 버전 정보

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.36.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.36.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.36.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.35.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-6.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-6.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.34.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.33.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.33.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.32.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.32.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.31.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.31.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.30.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.30.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.29.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.28.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.28.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.27.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.27.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.26.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.25.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.24.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.24.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.23.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.23.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.22.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.21.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.21.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.20.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.20.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.19.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.19.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.18.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.18.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.17.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.17.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.16.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.16.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.15.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.15.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.14.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.14.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.13.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.13.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.12.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.12.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.11.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.11.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.11.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.10.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.10.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.9.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, webserver
emr-5.5.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.5.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.4.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.3.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.2.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.2.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.1.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.1.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-5.0.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-5.0.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.9.6	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.9.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.9.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.8.5	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.8.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.8.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.7.4	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.3	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.2	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.7.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.7.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.6.1	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.6.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.5.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.4.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Amazon EMR 릴리스 레이블	Ganglia 버전	Ganglia와 함께 설치된 구성 요소
emr-4.3.0	3.7.2	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver
emr-4.2.0	3.6.0	emrfs, emr-goodies, ganglia-monitor, ganglia-metadata-collector, ganglia-web, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, webserver

Apache 하둡

Apache Hadoop은 오픈 소스 Java 소프트웨어 프레임워크로 인스턴스 클러스터 전체에 걸친 대량 데이터 처리를 지원합니다. 단일 인스턴스에서 실행될 수도 있고 수천 대의 인스턴스에서 실행될 수도 있습니다. 하둡은 MapReduce 및 Tez와 같은 많은 처리 모델을 사용하여 여러 인스턴스에 처리를 분산하고 HDFS라는 분산 파일 시스템을 사용하여 여러 인스턴스에 데이터를 저장합니다. 하둡은 클러스터에 속한 인스턴스의 상태를 모니터링하고, 하나 이상의 노드에서 장애가 발생할 경우 이를 복구할 수 있습니다. 이처럼 Hadoop은 처리 및 스토리지 용량을 증가시켜 줄 뿐 아니라 고가용성까지 보장합니다. 자세한 내용은 Hadoop 설명서를 참조하세요.

다음 테이블에는 Amazon EMR이 Hadoop을 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Hadoop의 버전이 나열되어 있습니다.

이 릴리스에서 하둡과 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 하둡 버전 정보

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-7.9.0	Hadoop 3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

다음 테이블에는 Amazon EMR이 Hadoop을 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Hadoop의 버전이 나열되어 있습니다.

이 릴리스에서 Hadoop과 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Hadoop 버전 정보

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.15.0	Hadoop 3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

다음 테이블에는 Amazon EMR이 Hadoop을 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Hadoop의 버전이 나열되어 있습니다.

이 릴리스에서 Hadoop과 함께 설치된 구성 요소의 버전은 [릴리스 5.36.2 구성 요소 버전]을 참조하세요.

emr-5.36.2용 Hadoop 버전 정보

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.36.2	Hadoop 2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resour

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
		cemanager, hadoop-yarn-timeline-server

Amazon EMR 5.18.0부터 Amazon EMR 아티팩트 리포지토리를 사용하여 특정 Amazon EMR 릴리스 버전에서 사용할 수 있는 정확한 버전의 라이브러리와 종속 항목에 대해 작업 코드를 빌드할 수 있습니다. 자세한 내용은 Amazon EMR 아티팩트 리포지토리를 사용하여 종속성 확인 단원을 참조하십시오.

주제

- 하둡 구성
- Amazon EMR에서 HDFS의 투명한 암호화
- Hadoop 애플리케이션 생성 또는 실행
- 복원된 객체 읽기
- YARN 컨테이너에 대해 비균일 메모리 액세스 인식 활성화
- YARN 컨테이너 빈 패킹
- Hadoop 버전 기록

하둡 구성

다음 단원에서는 Hadoop 데몬, 작업 및 HDFS의 기본 구성 설정을 제공합니다.

주제

- 작업 구성
- Hadoop 대몬(daemon) 구성 설정
- HDFS 구성

작업 구성

구성 변수를 설정하여 MapReduce 작업의 성능을 조정할 수 있습니다. 이 단원에는 중요한 설정의 기본값이 나와 있습니다. 기본값은 클러스터에서 사용되는 노드의 EC2 인스턴스 유형에 따라 달라집니다. Amazon EMR 릴리스 버전 4.6.0 이상을 사용하면 HBase를 사용할 수 있습니다. HBase가 설치된 경우 다른 기본값이 사용됩니다. 이러한 값은 초기 기본값과 함께 제공됩니다.

하둡 2는 `mapreduce.map.java.opts` 파라미터와 `mapreduce.reduce.java.opts` 파라미터를 사용하여 맵 및 JVM의 메모리를 각각 구성합니다. 이 두 파라미터는 이전 하둡 버전의 단일 `mapreduce.map.java.opts` 구성 옵션을 대신합니다.

마찬가지로, Hadoop 2.7.2 이상에서 `mapred.job.jvm.num.tasks`는 `mapred.job.reuse.jvm.num.tasks`를 대체합니다. Amazon EMR은 EC2 인스턴스 유형에 관계없이 이 값을 20으로 설정합니다. `mapred-site` 구성 분류를 사용하여 이 설정을 재정의할 수 있습니다. -1 값을 설정하면 단일 작업 내의 수많은 작업에 JVM을 다시 사용할 수 있으며 1 값을 설정하면 각 작업의 새 JVM이 생성됩니다.

예를 들어, `mapred.job.jvm.num.tasks` 값을 -1로 설정하면 다음 내용이 포함된 파일을 생성할 수 있습니다.

```
[
 {
  "Classification": "mapred-site",
  "Properties": {
   "mapred.job.jvm.num.tasks": "-1"
  }
 }
]
```

의 `create-cluster` 명령 또는 `modify-instance-groups` 명령을 사용하는 경우 JSON 구성 파일을 참조 AWS CLI할 수 있습니다. 다음 예제에서 구성 파일은 `myConfig.json`으로 Amazon S3에 저장됩니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge \
--instance-count 3 --applications Name=Hadoop --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json \
--use-default-roles
```

동일한 방법으로 mapred-site 구성 분류를 사용하여 아래 나열된 기본값을 변경할 수 있으며, 단일 JSON 파일을 사용하여 여러 값과 여러 구성 분류를 설정할 수 있습니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

작업 구성 설정의 기본값

> **Note**
>
> 각 인스턴스에 지정된 값은 최신 EMR 릴리스를 따릅니다. 이전 릴리스의 기본 설정은 다를 수 있습니다.

인스턴스 유형

- c1 인스턴스
- c3 인스턴스
- c4 인스턴스
- c5 인스턴스
- c5a 인스턴스
- c5ad 인스턴스
- c5d 인스턴스
- c5n 인스턴스
- c6a 인스턴스
- c6g 인스턴스
- c6gd 인스턴스
- c6gn 인스턴스
- c6i 인스턴스
- c6id 인스턴스
- c6in 인스턴스
- c7a 인스턴스

- c7g 인스턴스
- c7gd 인스턴스
- c7gn 인스턴스
- c7i 인스턴스
- c7i-flex 인스턴스
- c8g 인스턴스
- c8gd 인스턴스
- d2 인스턴스
- d3 인스턴스
- d3en 인스턴스
- f2 인스턴스
- g3 인스턴스
- g3s 인스턴스
- g4dn 인스턴스
- g5 인스턴스
- g6 인스턴스
- g6e 인스턴스
- gr6 인스턴스
- h1 인스턴스
- i2 인스턴스
- i3 인스턴스
- i3en 인스턴스
- i4g 인스턴스
- i4i 인스턴스
- i7i 인스턴스
- i7ie 인스턴스
- i8g 인스턴스
- im4gn 인스턴스
- is4gen 인스턴스

- m1 인스턴스
- m2 인스턴스
- m3 인스턴스
- m4 인스턴스
- m5 인스턴스
- m5a 인스턴스
- m5ad 인스턴스
- m5d 인스턴스
- m5dn 인스턴스
- m5n 인스턴스
- m5zn 인스턴스
- m6a 인스턴스
- m6g 인스턴스
- m6gd 인스턴스
- m6i 인스턴스
- m6id 인스턴스
- m6idn 인스턴스
- m6in 인스턴스
- m7a 인스턴스
- m7g 인스턴스
- m7gd 인스턴스
- m7i 인스턴스
- m7i-flex 인스턴스
- m8g 인스턴스
- m8gd 인스턴스
- p2 인스턴스
- p3 인스턴스
- p4d 인스턴스
- p5 인스턴스
- r3 인스턴스

- r4 인스턴스
- r5 인스턴스
- r5a 인스턴스
- r5ad 인스턴스
- r5b 인스턴스
- r5d 인스턴스
- r5dn 인스턴스
- r5n 인스턴스
- r6a 인스턴스
- r6g 인스턴스
- r6gd 인스턴스
- r6i 인스턴스
- r6id 인스턴스
- r6idn 인스턴스
- r6in 인스턴스
- r7a 인스턴스
- r7g 인스턴스
- r7gd 인스턴스
- r7i 인스턴스
- r7iz 인스턴스
- r8g 인스턴스
- r8gd 인스턴스
- x1 인스턴스
- x1e 인스턴스
- x2gd 인스턴스
- x2idn 인스턴스
- x2iedn 인스턴스
- x8g 인스턴스
- z1d 인스턴스

c1 인스턴스

c1.medium

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx288m	-Xmx288m
mapreduce.java.opts	-Xmx288m	-Xmx288m
mapreduce.map.memory.mb	512	512
mapreduce.reduce.memory.mb	512	512
yarn.app.mapreduce.am.resource.mb	512	512
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	512	512
yarn.nodemanager.resource.memory-mb	1024	512

c1.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1024	1024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	2048	2560
`yarn.nodemanager.resource.memory-mb`	5120	2560

c3 인스턴스

c3.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1126m	-Xmx1126m
`mapreduce.java.opts`	-Xmx2252m	-Xmx2252m
`mapreduce.map.memory.mb`	1408	1408
`mapreduce.reduce.memory.mb`	2816	2816
`yarn.app.mapreduce.am.resource.mb`	2816	2816
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	5632	2816
`yarn.nodemanager.resource.memory-mb`	5632	2816

c3.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2,880	2,880
yarn.app.mapreduce.am.resource.mb	2,880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

c3.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2,880	2,880
yarn.app.mapreduce.am.resource.mb	2,880	2880

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c3.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1331m	-Xmx1331m
mapreduce.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.map.memory.mb	1664	1664
mapreduce.reduce.memory.mb	3328	3328
yarn.app.mapreduce.am.resource.mb	3328	3328
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

c4 인스턴스

c4.large

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx717m	-Xmx717m
`mapreduce.java.opts`	-Xmx1434m	-Xmx1434m
`mapreduce.map.memory.mb`	896	896
`mapreduce.reduce.memory.mb`	1792	1792
`yarn.app.mapreduce.am.resource.mb`	1792	1792
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	1792	896
`yarn.nodemanager.resource.memory-mb`	1792	896

c4.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1126m	-Xmx1126m
`mapreduce.java.opts`	-Xmx2252m	-Xmx2252m
`mapreduce.map.memory.mb`	1408	1408
`mapreduce.reduce.memory.mb`	2816	2816
`yarn.app.mapreduce.am.resource.mb`	2816	2816

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c4.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2,880	2,880
yarn.app.mapreduce.am.resource.mb	2,880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

c4.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2,880	2,880
yarn.app.mapreduce.am.resource.mb	2,880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

c4.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1183m	-Xmx1183m
mapreduce.java.opts	-Xmx2366m	-Xmx2366m
mapreduce.map.memory.mb	1479	1479
mapreduce.reduce.memory.mb	2958	2958
yarn.app.mapreduce.am.resource.mb	2958	2958

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

c5 인스턴스

c5.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	6144	3072
yarn.nodemanager.resource.memory-mb	6144	3072

c5.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5.9xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1502m	-Xmx1502m
`mapreduce.java.opts`	-Xmx3004m	-Xmx3004m
`mapreduce.map.memory.mb`	1877	1877
`mapreduce.reduce.memory.mb`	3754	3754
`yarn.app.mapreduce.am.resource.mb`	3754	3754
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	90112	30048
`yarn.nodemanager.resource.memory-mb`	90112	30048

c5.18xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1547m	-Xmx1547m
`mapreduce.java.opts`	-Xmx3094m	-Xmx3094m
`mapreduce.map.memory.mb`	1934	1934
`mapreduce.reduce.memory.mb`	3868	3868
`yarn.app.mapreduce.am.resource.mb`	3868	3868

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

c5a 인스턴스

c5a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1126m	-Xmx1126m
`mapreduce.java.opts`	-Xmx2252m	-Xmx2252m
`mapreduce.map.memory.mb`	1408	1408
`mapreduce.reduce.memory.mb`	2816	2816
`yarn.app.mapreduce.am.resource.mb`	2816	2816
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	5632	2816
`yarn.nodemanager.resource.memory-mb`	5632	2816

c5a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c5a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1357m	-Xmx1357m
`mapreduce.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.map.memory.mb`	1696	1696
`mapreduce.reduce.memory.mb`	3392	3392
`yarn.app.mapreduce.am.resource.mb`	3392	3392
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

c5a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1502m	-Xmx1502m
`mapreduce.java.opts`	-Xmx3004m	-Xmx3004m
`mapreduce.map.memory.mb`	1877	1877
`mapreduce.reduce.memory.mb`	3754	3754
`yarn.app.mapreduce.am.resource.mb`	3754	3754

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30048
yarn.nodemanager.resource.memory-mb	90112	30048

c5a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1494m	-Xmx1494m
`mapreduce.java.opts`	-Xmx2988m	-Xmx2988m
`mapreduce.map.memory.mb`	1867	1867
`mapreduce.reduce.memory.mb`	3734	3734
`yarn.app.mapreduce.am.resource.mb`	3734	3734
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	179200	29840
`yarn.nodemanager.resource.memory-mb`	179200	29840

c5ad 인스턴스

c5ad.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1126m	-Xmx1126m
`mapreduce.java.opts`	-Xmx2252m	-Xmx2252m
`mapreduce.map.memory.mb`	1408	1408
`mapreduce.reduce.memory.mb`	2816	2816
`yarn.app.mapreduce.am.resource.mb`	2816	2816

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c5ad.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c5ad.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

c5ad.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1357m	-Xmx1357m
`mapreduce.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.map.memory.mb`	1696	1696
`mapreduce.reduce.memory.mb`	3392	3392
`yarn.app.mapreduce.am.resource.mb`	3392	3392

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c5ad.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c5ad.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c5ad.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	179200	29840
`yarn.nodemanager.resource.memory-mb`	179200	29840

c5d 인스턴스

c5d.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1229m	-Xmx1229m
`mapreduce.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.map.memory.mb`	1536	1536
`mapreduce.reduce.memory.mb`	3072	3072
`yarn.app.mapreduce.am.resource.mb`	3072	3072
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	6144	3072
`yarn.nodemanager.resource.memory-mb`	6144	3072

c5d.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

c5d.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1229m	-Xmx1229m
mapreduce.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	3072	3072
yarn.app.mapreduce.am.resource.mb	3072	3072

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

c5d.9xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1456m	-Xmx1456m
mapreduce.java.opts	-Xmx2912m	-Xmx2912m
mapreduce.map.memory.mb	1820	1820
mapreduce.reduce.memory.mb	3640	3640
yarn.app.mapreduce.am.resource.mb	3640	3640
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	65536	32768
yarn.nodemanager.resource.memory-mb	65536	32768

c5d.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1502m	-Xmx1502m
`mapreduce.java.opts`	-Xmx3004m	-Xmx3004m
`mapreduce.map.memory.mb`	1877	1877
`mapreduce.reduce.memory.mb`	3754	3754
`yarn.app.mapreduce.am.resource.mb`	3754	3754
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	90112	30048
`yarn.nodemanager.resource.memory-mb`	90112	30048

c5d.18xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1547m	-Xmx1547m
`mapreduce.java.opts`	-Xmx3094m	-Xmx3094m
`mapreduce.map.memory.mb`	1934	1934
`mapreduce.reduce.memory.mb`	3868	3868
`yarn.app.mapreduce.am.resource.mb`	3868	3868

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	139264	30960
yarn.nodemanager.resource.memory-mb	139264	30960

c5d.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1570m	-Xmx1570m
mapreduce.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.map.memory.mb	1963	1963
mapreduce.reduce.memory.mb	3926	3926
yarn.app.mapreduce.am.resource.mb	3926	3926
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31376
yarn.nodemanager.resource.memory-mb	188416	31376

c5n 인스턴스

c5n.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1613m	-Xmx1613m
`mapreduce.java.opts`	-Xmx3226m	-Xmx3226m
`mapreduce.map.memory.mb`	2016	2016
`mapreduce.reduce.memory.mb`	4032	4032
`yarn.app.mapreduce.am.resource.mb`	4032	4032
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	8064	4032
`yarn.nodemanager.resource.memory-mb`	8064	4032

c5n.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1613m	-Xmx1613m
`mapreduce.java.opts`	-Xmx3226m	-Xmx3226m
`mapreduce.map.memory.mb`	2016	2016
`mapreduce.reduce.memory.mb`	4032	4032
`yarn.app.mapreduce.am.resource.mb`	4032	4032

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	16128	8064
`yarn.nodemanager.resource.memory-mb`	16128	8064

c5n.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1741m	-Xmx1741m
`mapreduce.java.opts`	-Xmx3482m	-Xmx3482m
`mapreduce.map.memory.mb`	2176	2176
`mapreduce.reduce.memory.mb`	4352	4352
`yarn.app.mapreduce.am.resource.mb`	4352	4352
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	34816	17408
`yarn.nodemanager.resource.memory-mb`	34816	17408

c5n.9xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2002m	-Xmx2002m
mapreduce.java.opts	-Xmx4004m	-Xmx4004m
mapreduce.map.memory.mb	2503	2503
mapreduce.reduce.memory.mb	5006	5006
yarn.app.mapreduce.am.resource.mb	5006	5006
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

c5n.18xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2094m	-Xmx2094m
mapreduce.java.opts	-Xmx4188m	-Xmx4188m
mapreduce.map.memory.mb	2617	2617
mapreduce.reduce.memory.mb	5234	5234
yarn.app.mapreduce.am.resource.mb	5234	5234

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

c6a 인스턴스

c6a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6a.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c6a.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

c6g 인스턴스

c6g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1425m	-Xmx1425m
`mapreduce.java.opts`	-Xmx2850m	-Xmx2850m
`mapreduce.map.memory.mb`	1781	1781
`mapreduce.reduce.memory.mb`	3562	3562
`yarn.app.mapreduce.am.resource.mb`	3562	3562
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	85504	32074
`yarn.nodemanager.resource.memory-mb`	85504	32074

c6g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1459m	-Xmx1459m
`mapreduce.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.map.memory.mb`	1824	1824
`mapreduce.reduce.memory.mb`	3648	3648
`yarn.app.mapreduce.am.resource.mb`	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6gd 인스턴스

c6gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

c6gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6gn 인스턴스

c6gn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6gn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6gn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6gn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6gn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6gn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6i 인스턴스

c6i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6i.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

c6i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6i.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c6id 인스턴스

c6id.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6id.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6id.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6id.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c6id.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6id.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6id.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6id.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c6in 인스턴스

c6in.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c6in.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c6in.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c6in.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1357m	-Xmx1357m
`mapreduce.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.map.memory.mb`	1696	1696
`mapreduce.reduce.memory.mb`	3392	3392
`yarn.app.mapreduce.am.resource.mb`	3392	3392
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

c6in.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1425m	-Xmx1425m
`mapreduce.java.opts`	-Xmx2850m	-Xmx2850m
`mapreduce.map.memory.mb`	1781	1781
`mapreduce.reduce.memory.mb`	3562	3562
`yarn.app.mapreduce.am.resource.mb`	3562	3562

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c6in.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c6in.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c6in.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c7a 인스턴스

c7a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7a.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1510m	-Xmx1510m
mapreduce.java.opts	-Xmx3020m	-Xmx3020m
mapreduce.map.memory.mb	1888	1888
mapreduce.reduce.memory.mb	3776	3776
yarn.app.mapreduce.am.resource.mb	3776	3776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

c7a.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

c7g 인스턴스

c7g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

c7g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	116736	29184
`yarn.nodemanager.resource.memory-mb`	116736	29184

c7gd 인스턴스

c7gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1126m	-Xmx1126m
`mapreduce.java.opts`	-Xmx2252m	-Xmx2252m
`mapreduce.map.memory.mb`	1408	1408
`mapreduce.reduce.memory.mb`	2816	2816
`yarn.app.mapreduce.am.resource.mb`	2816	2816
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	5632	2816
`yarn.nodemanager.resource.memory-mb`	5632	2816

c7gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1425m	-Xmx1425m
`mapreduce.java.opts`	-Xmx2850m	-Xmx2850m
`mapreduce.map.memory.mb`	1781	1781
`mapreduce.reduce.memory.mb`	3562	3562
`yarn.app.mapreduce.am.resource.mb`	3562	3562
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	85504	32074
`yarn.nodemanager.resource.memory-mb`	85504	32074

c7gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1459m	-Xmx1459m
`mapreduce.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.map.memory.mb`	1824	1824
`mapreduce.reduce.memory.mb`	3648	3648
`yarn.app.mapreduce.am.resource.mb`	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7gn 인스턴스

c7gn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7gn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

c7gn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7gn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7gn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7gn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7i 인스턴스

c7i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c7i.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c7i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1425m	-Xmx1425m
`mapreduce.java.opts`	-Xmx2850m	-Xmx2850m
`mapreduce.map.memory.mb`	1781	1781
`mapreduce.reduce.memory.mb`	3562	3562
`yarn.app.mapreduce.am.resource.mb`	3562	3562
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	85504	32074
`yarn.nodemanager.resource.memory-mb`	85504	32074

c7i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1459m	-Xmx1459m
`mapreduce.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.map.memory.mb`	1824	1824
`mapreduce.reduce.memory.mb`	3648	3648
`yarn.app.mapreduce.am.resource.mb`	3648	3648

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c7i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c7i.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

c7i-flex 인스턴스

c7i-flex.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	5632	2816
`yarn.nodemanager.resource.memory-mb`	5632	2816

c7i-flex.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

c7i-flex.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c7i-flex.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c7i-flex.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c7i-flex.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8g 인스턴스

c8g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c8g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

c8g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

c8g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1357m	-Xmx1357m
mapreduce.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.map.memory.mb	1696	1696
mapreduce.reduce.memory.mb	3392	3392
yarn.app.mapreduce.am.resource.mb	3392	3392

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

c8g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1425m	-Xmx1425m
mapreduce.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.map.memory.mb	1781	1781
mapreduce.reduce.memory.mb	3562	3562
yarn.app.mapreduce.am.resource.mb	3562	3562
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c8g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8g.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1494m	-Xmx1494m
mapreduce.java.opts	-Xmx2988m	-Xmx2988m
mapreduce.map.memory.mb	1867	1867
mapreduce.reduce.memory.mb	3734	3734
yarn.app.mapreduce.am.resource.mb	3734	3734

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29840
yarn.nodemanager.resource.memory-mb	179200	29840

c8g.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1527m	-Xmx1527m
mapreduce.java.opts	-Xmx3054m	-Xmx3054m
mapreduce.map.memory.mb	1909	1909
mapreduce.reduce.memory.mb	3818	3818
yarn.app.mapreduce.am.resource.mb	3818	3818
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

c8gd 인스턴스

c8gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1126m	-Xmx1126m
mapreduce.java.opts	-Xmx2252m	-Xmx2252m
mapreduce.map.memory.mb	1408	1408
mapreduce.reduce.memory.mb	2816	2816
yarn.app.mapreduce.am.resource.mb	2816	2816
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	5632	2816
yarn.nodemanager.resource.memory-mb	5632	2816

c8gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1171m	-Xmx1171m
mapreduce.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.map.memory.mb	1464	1464
mapreduce.reduce.memory.mb	2928	2928
yarn.app.mapreduce.am.resource.mb	2928	2928

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

c8gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1171m	-Xmx1171m
`mapreduce.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.map.memory.mb`	1464	1464
`mapreduce.reduce.memory.mb`	2928	2928
`yarn.app.mapreduce.am.resource.mb`	2928	2928
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

c8gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1357m	-Xmx1357m
`mapreduce.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.map.memory.mb`	1696	1696
`mapreduce.reduce.memory.mb`	3392	3392
`yarn.app.mapreduce.am.resource.mb`	3392	3392
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

c8gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1425m	-Xmx1425m
`mapreduce.java.opts`	-Xmx2850m	-Xmx2850m
`mapreduce.map.memory.mb`	1781	1781
`mapreduce.reduce.memory.mb`	3562	3562
`yarn.app.mapreduce.am.resource.mb`	3562	3562

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32074
yarn.nodemanager.resource.memory-mb	85504	32074

c8gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1459m	-Xmx1459m
mapreduce.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.map.memory.mb	1824	1824
mapreduce.reduce.memory.mb	3648	3648
yarn.app.mapreduce.am.resource.mb	3648	3648
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

c8gd.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1494m	-Xmx1494m
`mapreduce.java.opts`	-Xmx2988m	-Xmx2988m
`mapreduce.map.memory.mb`	1867	1867
`mapreduce.reduce.memory.mb`	3734	3734
`yarn.app.mapreduce.am.resource.mb`	3734	3734
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	179200	29840
`yarn.nodemanager.resource.memory-mb`	179200	29840

c8gd.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1527m	-Xmx1527m
`mapreduce.java.opts`	-Xmx3054m	-Xmx3054m
`mapreduce.map.memory.mb`	1909	1909
`mapreduce.reduce.memory.mb`	3818	3818
`yarn.app.mapreduce.am.resource.mb`	3818	3818

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30608
yarn.nodemanager.resource.memory-mb	366592	30608

d2 인스턴스

d2.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d2.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

d2.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.java.opts`	-Xmx5836m	-Xmx5836m
`mapreduce.map.memory.mb`	3648	3648
`mapreduce.reduce.memory.mb`	7296	7296
`yarn.app.mapreduce.am.resource.mb`	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d2.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2417m	-Xmx2417m
mapreduce.java.opts	-Xmx4834m	-Xmx4834m
mapreduce.map.memory.mb	3021	3021
mapreduce.reduce.memory.mb	6042	6042
yarn.app.mapreduce.am.resource.mb	6042	6042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30194
yarn.nodemanager.resource.memory-mb	241664	30194

d3 인스턴스

d3.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.java.opts`	-Xmx12084m	-Xmx12084m
`mapreduce.map.memory.mb`	7552	7552
`mapreduce.reduce.memory.mb`	15104	15104
`yarn.app.mapreduce.am.resource.mb`	15104	15104
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

d3en 인스턴스

d3en.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

d3en.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

d3en.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

d3en.6xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

d3en.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

d3en.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

f2 인스턴스

f2.6xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx8055m	-Xmx8055m
mapreduce.java.opts	-Xmx16110m	-Xmx16110m
mapreduce.map.memory.mb	10069	10069
mapreduce.reduce.memory.mb	20138	20138
yarn.app.mapreduce.am.resource.mb	20138	20138

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	20146
`yarn.nodemanager.resource.memory-mb`	241664	20146

f2.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx8192m	-Xmx8192m
`mapreduce.java.opts`	-Xmx16384m	-Xmx16384m
`mapreduce.map.memory.mb`	10240	10240
`mapreduce.reduce.memory.mb`	20480	20480
`yarn.app.mapreduce.am.resource.mb`	20480	20480
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	491520	20480
`yarn.nodemanager.resource.memory-mb`	491520	20480

f2.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx8294m	-Xmx8294m
`mapreduce.java.opts`	-Xmx16588m	-Xmx16588m
`mapreduce.map.memory.mb`	10368	10368
`mapreduce.reduce.memory.mb`	20736	20736
`yarn.app.mapreduce.am.resource.mb`	20736	20736
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	1990656	20736
`yarn.nodemanager.resource.memory-mb`	1990656	20736

g3 인스턴스

g3.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g3.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g3.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

g3s 인스턴스

g3s.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g4dn 인스턴스

g4dn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

g4dn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

g4dn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

g4dn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

g4dn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

g4dn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

g5 인스턴스

g5.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

g5.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

g5.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g5.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g5.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

g5.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g5.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

g5.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

g6 인스턴스

g6.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2,880	2880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11520	5760
`yarn.nodemanager.resource.memory-mb`	11520	5760

g6.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2304m	-Xmx2304m
`mapreduce.java.opts`	-Xmx4608m	-Xmx4608m
`mapreduce.map.memory.mb`	2,880	2880
`mapreduce.reduce.memory.mb`	5760	5760
`yarn.app.mapreduce.am.resource.mb`	5760	5760
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23040	11520
`yarn.nodemanager.resource.memory-mb`	23040	11520

g6.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2662m	-Xmx2662m
mapreduce.java.opts	-Xmx5324m	-Xmx5324m
mapreduce.map.memory.mb	3328	3328
mapreduce.reduce.memory.mb	6656	6656
yarn.app.mapreduce.am.resource.mb	6656	6656
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	53248	26624
yarn.nodemanager.resource.memory-mb	53248	26624

g6.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2893m	-Xmx2893m
mapreduce.java.opts	-Xmx5786m	-Xmx5786m
mapreduce.map.memory.mb	3616	3616
mapreduce.reduce.memory.mb	7232	7232
yarn.app.mapreduce.am.resource.mb	7232	7232

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

g6.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2953m	-Xmx2953m
mapreduce.java.opts	-Xmx5906m	-Xmx5906m
mapreduce.map.memory.mb	3691	3691
mapreduce.reduce.memory.mb	7382	7382
yarn.app.mapreduce.am.resource.mb	7382	7382
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	177152	29512
yarn.nodemanager.resource.memory-mb	177152	29512

g6.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2995m	-Xmx2995m
mapreduce.java.opts	-Xmx5990m	-Xmx5990m
mapreduce.map.memory.mb	3744	3744
mapreduce.reduce.memory.mb	7488	7488
yarn.app.mapreduce.am.resource.mb	7488	7488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	239616	29952
yarn.nodemanager.resource.memory-mb	239616	29952

g6.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3038m	-Xmx3038m
mapreduce.java.opts	-Xmx6076m	-Xmx6076m
mapreduce.map.memory.mb	3797	3797
mapreduce.reduce.memory.mb	7594	7594
yarn.app.mapreduce.am.resource.mb	7594	7594

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	364544	30408
yarn.nodemanager.resource.memory-mb	364544	30408

g6.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	737280	30720
yarn.nodemanager.resource.memory-mb	737280	30720

g6e 인스턴스

g6e.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

g6e.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

g6e.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

g6e.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

g6e.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

g6e.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

g6e.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

g6e.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

gr6 인스턴스

gr6.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5786m	-Xmx5786m
mapreduce.java.opts	-Xmx11572m	-Xmx11572m
mapreduce.map.memory.mb	7232	7232
mapreduce.reduce.memory.mb	14464	14464
yarn.app.mapreduce.am.resource.mb	14464	14464
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	115712	28928
yarn.nodemanager.resource.memory-mb	115712	28928

gr6.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5990m	-Xmx5990m
`mapreduce.java.opts`	-Xmx11980m	-Xmx11980m
`mapreduce.map.memory.mb`	7488	7488
`mapreduce.reduce.memory.mb`	14976	14976
`yarn.app.mapreduce.am.resource.mb`	14976	14976
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	239616	29952
`yarn.nodemanager.resource.memory-mb`	239616	29952

h1 인스턴스

h1.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.java.opts`	-Xmx4916m	-Xmx4916m
`mapreduce.map.memory.mb`	3072	3072
`mapreduce.reduce.memory.mb`	6144	6144
`yarn.app.mapreduce.am.resource.mb`	6144	6144

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

h1.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

h1.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3072m	-Xmx3072m
`mapreduce.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.map.memory.mb`	3840	3840
`mapreduce.reduce.memory.mb`	7680	7680
`yarn.app.mapreduce.am.resource.mb`	7680	7680
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	122880	30720
`yarn.nodemanager.resource.memory-mb`	122880	30720

h1.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3174m	-Xmx3174m
`mapreduce.java.opts`	-Xmx6348m	-Xmx6348m
`mapreduce.map.memory.mb`	3968	3968
`mapreduce.reduce.memory.mb`	7936	7936
`yarn.app.mapreduce.am.resource.mb`	7936	7936

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

i2 인스턴스

i2.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i2.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i2.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i2.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i3 인스턴스

i3.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i3.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i3.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i3.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i3.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i3en 인스턴스

i3en.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

i3en.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

i3en.3xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

i3en.6xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

i3en.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

i3en.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

i4g 인스턴스

i4g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5427m	-Xmx5427m
`mapreduce.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.map.memory.mb`	6784	6784
`mapreduce.reduce.memory.mb`	13568	13568
`yarn.app.mapreduce.am.resource.mb`	13568	13568
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

i4g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i4g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i4i 인스턴스

i4i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i4i.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i4i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	116736	29184
`yarn.nodemanager.resource.memory-mb`	116736	29184

i4i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.java.opts`	-Xmx12084m	-Xmx12084m
`mapreduce.map.memory.mb`	7552	7552
`mapreduce.reduce.memory.mb`	15104	15104
`yarn.app.mapreduce.am.resource.mb`	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

i4i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i4i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i4i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i4i.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

i7i 인스턴스

i7i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7i.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i7i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i7i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.java.opts`	-Xmx12084m	-Xmx12084m
`mapreduce.map.memory.mb`	7552	7552
`mapreduce.reduce.memory.mb`	15104	15104
`yarn.app.mapreduce.am.resource.mb`	15104	15104
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

i7i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.java.opts`	-Xmx12220m	-Xmx12220m
`mapreduce.map.memory.mb`	7637	7637
`mapreduce.reduce.memory.mb`	15274	15274
`yarn.app.mapreduce.am.resource.mb`	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

i7i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7i.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

i7ie 인스턴스

i7ie.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i7ie.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5427m	-Xmx5427m
`mapreduce.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.map.memory.mb`	6784	6784
`mapreduce.reduce.memory.mb`	13568	13568
`yarn.app.mapreduce.am.resource.mb`	13568	13568
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

i7ie.3xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5700m	-Xmx5700m
`mapreduce.java.opts`	-Xmx11400m	-Xmx11400m
`mapreduce.map.memory.mb`	7125	7125
`mapreduce.reduce.memory.mb`	14250	14250
`yarn.app.mapreduce.am.resource.mb`	14250	14250

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28504
yarn.nodemanager.resource.memory-mb	85504	28504

i7ie.6xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5974m	-Xmx5974m
mapreduce.java.opts	-Xmx11948m	-Xmx11948m
mapreduce.map.memory.mb	7467	7467
mapreduce.reduce.memory.mb	14934	14934
yarn.app.mapreduce.am.resource.mb	14934	14934
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29860
yarn.nodemanager.resource.memory-mb	179200	29860

i7ie.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

i7ie.18xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6155m	-Xmx6155m
mapreduce.java.opts	-Xmx12310m	-Xmx12310m
mapreduce.map.memory.mb	7694	7694
mapreduce.reduce.memory.mb	15388	15388
yarn.app.mapreduce.am.resource.mb	15388	15388

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	553984	30792
yarn.nodemanager.resource.memory-mb	553984	30792

i7ie.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i7ie.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6212m	-Xmx6212m
`mapreduce.java.opts`	-Xmx12424m	-Xmx12424m
`mapreduce.map.memory.mb`	7765	7765
`mapreduce.reduce.memory.mb`	15530	15530
`yarn.app.mapreduce.am.resource.mb`	15530	15530
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	1490944	31124
`yarn.nodemanager.resource.memory-mb`	1490944	31124

i8g 인스턴스

i8g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx4685m	-Xmx4685m
`mapreduce.java.opts`	-Xmx9370m	-Xmx9370m
`mapreduce.map.memory.mb`	5856	5856
`mapreduce.reduce.memory.mb`	11712	11712
`yarn.app.mapreduce.am.resource.mb`	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

i8g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

i8g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

i8g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

i8g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.java.opts`	-Xmx12220m	-Xmx12220m
`mapreduce.map.memory.mb`	7637	7637
`mapreduce.reduce.memory.mb`	15274	15274
`yarn.app.mapreduce.am.resource.mb`	15274	15274
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30564
`yarn.nodemanager.resource.memory-mb`	366592	30564

i8g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.java.opts`	-Xmx12288m	-Xmx12288m
`mapreduce.map.memory.mb`	7680	7680
`mapreduce.reduce.memory.mb`	15360	15360
`yarn.app.mapreduce.am.resource.mb`	15360	15360
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	491520	30720
`yarn.nodemanager.resource.memory-mb`	491520	30720

i8g.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6178m	-Xmx6178m
`mapreduce.java.opts`	-Xmx12356m	-Xmx12356m
`mapreduce.map.memory.mb`	7723	7723
`mapreduce.reduce.memory.mb`	15446	15446
`yarn.app.mapreduce.am.resource.mb`	15446	15446

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

i8g.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

im4gn 인스턴스

im4gn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

im4gn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

im4gn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

im4gn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

im4gn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

is4gen 인스턴스

is4gen.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3514m	-Xmx3514m
mapreduce.java.opts	-Xmx7028m	-Xmx7028m
mapreduce.map.memory.mb	4393	4393
mapreduce.reduce.memory.mb	8786	8786
yarn.app.mapreduce.am.resource.mb	8786	8786
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	17572.12	8786.06
yarn.nodemanager.resource.memory-mb	17572.12	8786.06

4gn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3866m	-Xmx3866m
`mapreduce.java.opts`	-Xmx7732m	-Xmx7732m
`mapreduce.map.memory.mb`	4832	4832
`mapreduce.reduce.memory.mb`	9664	9664
`yarn.app.mapreduce.am.resource.mb`	9664	9664
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	38656	19328
`yarn.nodemanager.resource.memory-mb`	38656	19328

is4gen.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx4275m	-Xmx4275m
`mapreduce.java.opts`	-Xmx8550m	-Xmx8550m
`mapreduce.map.memory.mb`	5344	5344
`mapreduce.reduce.memory.mb`	10688	10688
`yarn.app.mapreduce.am.resource.mb`	10688	10688

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	32064
yarn.nodemanager.resource.memory-mb	85504	32064

is4gen.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4480m	-Xmx4480m
mapreduce.java.opts	-Xmx8960m	-Xmx8960m
mapreduce.map.memory.mb	5600	5600
mapreduce.reduce.memory.mb	11200	11200
yarn.app.mapreduce.am.resource.mb	11200	11200
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	22400
yarn.nodemanager.resource.memory-mb	179200	22400

m1 인스턴스

m1.small

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx288m	-Xmx288m
`mapreduce.java.opts`	-Xmx288m	-Xmx288m
`mapreduce.map.memory.mb`	512	512
`mapreduce.reduce.memory.mb`	512	512
`yarn.app.mapreduce.am.resource.mb`	512	512
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	512	512
`yarn.nodemanager.resource.memory-mb`	1024	512

m1.medium

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx512m	-Xmx512m
`mapreduce.java.opts`	-Xmx768m	-Xmx768m
`mapreduce.map.memory.mb`	768	768
`mapreduce.reduce.memory.mb`	1024	1024
`yarn.app.mapreduce.am.resource.mb`	1024	1024

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	2048	1024
`yarn.nodemanager.resource.memory-mb`	2048	1024

m1.large

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx512m	-Xmx512m
`mapreduce.java.opts`	-Xmx1024m	-Xmx1024m
`mapreduce.map.memory.mb`	768	768
`mapreduce.reduce.memory.mb`	1536	1536
`yarn.app.mapreduce.am.resource.mb`	1536	1536
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	3072	2560
`yarn.nodemanager.resource.memory-mb`	5120	2560

m1.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx512m	-Xmx512m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	768	768
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m2 인스턴스

m2.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx864m	-Xmx864m
mapreduce.java.opts	-Xmx1536m	-Xmx1536m
mapreduce.map.memory.mb	1024	1024
mapreduce.reduce.memory.mb	2048	2048
yarn.app.mapreduce.am.resource.mb	2048	2048

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	7168	7168
yarn.nodemanager.resource.memory-mb	14336	7168

m2.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	15360
yarn.nodemanager.resource.memory-mb	30720	15360

m2.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1280m	-Xmx1280m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1536	1536
mapreduce.reduce.memory.mb	2560	2560
yarn.app.mapreduce.am.resource.mb	2560	2560
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	8192	30720
yarn.nodemanager.resource.memory-mb	61440	30720

m3 인스턴스

m3.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2,880	2,880
yarn.app.mapreduce.am.resource.mb	2,880	2880

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m3.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1152m	-Xmx1152m
mapreduce.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.map.memory.mb	1440	1440
mapreduce.reduce.memory.mb	2,880	2,880
yarn.app.mapreduce.am.resource.mb	2,880	2880
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23040	11520
yarn.nodemanager.resource.memory-mb	23040	11520

m4 인스턴스

m4.large

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1229m	-Xmx1229m
`mapreduce.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.map.memory.mb`	1536	1536
`mapreduce.reduce.memory.mb`	3072	3072
`yarn.app.mapreduce.am.resource.mb`	3072	3072
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	6144	3072
`yarn.nodemanager.resource.memory-mb`	6144	3072

m4.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1229m	-Xmx1229m
`mapreduce.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.map.memory.mb`	1536	1536
`mapreduce.reduce.memory.mb`	3072	3072
`yarn.app.mapreduce.am.resource.mb`	3072	3072

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	12288	6144
`yarn.nodemanager.resource.memory-mb`	12288	6144

m4.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1229m	-Xmx1229m
`mapreduce.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.map.memory.mb`	1536	1536
`mapreduce.reduce.memory.mb`	3072	3072
`yarn.app.mapreduce.am.resource.mb`	3072	3072
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	24576	12288
`yarn.nodemanager.resource.memory-mb`	24576	12288

m4.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1434m	-Xmx1434m
`mapreduce.java.opts`	-Xmx2868m	-Xmx2868m
`mapreduce.map.memory.mb`	1792	1792
`mapreduce.reduce.memory.mb`	3584	3584
`yarn.app.mapreduce.am.resource.mb`	3584	3584
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	57344	28672
`yarn.nodemanager.resource.memory-mb`	57344	28672

m4.10xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx1557m	-Xmx1557m
`mapreduce.java.opts`	-Xmx3114m	-Xmx3114m
`mapreduce.map.memory.mb`	1946	1946
`mapreduce.reduce.memory.mb`	3892	3892
`yarn.app.mapreduce.am.resource.mb`	3892	3892

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	155648	31104
yarn.nodemanager.resource.memory-mb	155648	31104

m4.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx1587m	-Xmx1587m
mapreduce.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.map.memory.mb	1984	1984
mapreduce.reduce.memory.mb	3968	3968
yarn.app.mapreduce.am.resource.mb	3968	3968
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5 인스턴스

m5.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.java.opts`	-Xmx4916m	-Xmx4916m
`mapreduce.map.memory.mb`	3072	3072
`mapreduce.reduce.memory.mb`	6144	6144
`yarn.app.mapreduce.am.resource.mb`	6144	6144
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	12288	6144
`yarn.nodemanager.resource.memory-mb`	12288	6144

m5.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.java.opts`	-Xmx4916m	-Xmx4916m
`mapreduce.map.memory.mb`	3072	3072
`mapreduce.reduce.memory.mb`	6144	6144
`yarn.app.mapreduce.am.resource.mb`	6144	6144

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3174m	-Xmx3174m
mapreduce.java.opts	-Xmx6348m	-Xmx6348m
mapreduce.map.memory.mb	3968	3968
mapreduce.reduce.memory.mb	7936	7936
yarn.app.mapreduce.am.resource.mb	7936	7936
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

m5.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

m5a 인스턴스

m5a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	12288	6144
yarn.nodemanager.resource.memory-mb	12288	6144

m5a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2458m	-Xmx2458m
mapreduce.java.opts	-Xmx4916m	-Xmx4916m
mapreduce.map.memory.mb	3072	3072
mapreduce.reduce.memory.mb	6144	6144
yarn.app.mapreduce.am.resource.mb	6144	6144
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31416
yarn.nodemanager.resource.memory-mb	188416	31416

m5a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3174m	-Xmx3174m
`mapreduce.java.opts`	-Xmx6348m	-Xmx6348m
`mapreduce.map.memory.mb`	3968	3968
`mapreduce.reduce.memory.mb`	7936	7936
`yarn.app.mapreduce.am.resource.mb`	7936	7936
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	253952	31744
`yarn.nodemanager.resource.memory-mb`	253952	31744

m5a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3209m	-Xmx3209m
`mapreduce.java.opts`	-Xmx6418m	-Xmx6418m
`mapreduce.map.memory.mb`	4011	4011
`mapreduce.reduce.memory.mb`	8022	8022
`yarn.app.mapreduce.am.resource.mb`	8022	8022

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

m5ad 인스턴스

m5ad.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5ad.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

m5ad.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

m5ad.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.java.opts`	-Xmx5836m	-Xmx5836m
`mapreduce.map.memory.mb`	3648	3648
`mapreduce.reduce.memory.mb`	7296	7296
`yarn.app.mapreduce.am.resource.mb`	7296	7296
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	116736	29184
`yarn.nodemanager.resource.memory-mb`	116736	29184

m5ad.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5ad.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5ad.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m5d 인스턴스

m5d.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.java.opts`	-Xmx4916m	-Xmx4916m
`mapreduce.map.memory.mb`	3072	3072
`mapreduce.reduce.memory.mb`	6144	6144
`yarn.app.mapreduce.am.resource.mb`	6144	6144
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	12288	6144
`yarn.nodemanager.resource.memory-mb`	12288	6144

m5d.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2458m	-Xmx2458m
`mapreduce.java.opts`	-Xmx4916m	-Xmx4916m
`mapreduce.map.memory.mb`	3072	3072
`mapreduce.reduce.memory.mb`	6144	6144
`yarn.app.mapreduce.am.resource.mb`	6144	6144

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

m5d.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2867m	-Xmx2867m
mapreduce.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.map.memory.mb	3584	3584
mapreduce.reduce.memory.mb	7168	7168
yarn.app.mapreduce.am.resource.mb	7168	7168
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

m5d.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

m5d.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3140m	-Xmx3140m
mapreduce.java.opts	-Xmx6280m	-Xmx6280m
mapreduce.map.memory.mb	3925	3925
mapreduce.reduce.memory.mb	7850	7850
yarn.app.mapreduce.am.resource.mb	7850	7850

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	188416	31416
`yarn.nodemanager.resource.memory-mb`	188416	31416

m5d.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3174m	-Xmx3174m
`mapreduce.java.opts`	-Xmx6348m	-Xmx6348m
`mapreduce.map.memory.mb`	3968	3968
`mapreduce.reduce.memory.mb`	7936	7936
`yarn.app.mapreduce.am.resource.mb`	7936	7936
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	253952	31744
`yarn.nodemanager.resource.memory-mb`	253952	31744

m5d.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3209m	-Xmx3209m
mapreduce.java.opts	-Xmx6418m	-Xmx6418m
mapreduce.map.memory.mb	4011	4011
mapreduce.reduce.memory.mb	8022	8022
yarn.app.mapreduce.am.resource.mb	8022	8022
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32056
yarn.nodemanager.resource.memory-mb	385024	32056

m5dn 인스턴스

m5dn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m5dn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5dn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5dn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5dn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m5dn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m5dn.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30520
`yarn.nodemanager.resource.memory-mb`	366592	30520

m5n 인스턴스

m5n.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

m5n.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

m5n.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m5n.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m5n.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2986m	-Xmx2986m
`mapreduce.java.opts`	-Xmx5972m	-Xmx5972m
`mapreduce.map.memory.mb`	3733	3733
`mapreduce.reduce.memory.mb`	7466	7466
`yarn.app.mapreduce.am.resource.mb`	7466	7466
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	179200	29880
`yarn.nodemanager.resource.memory-mb`	179200	29880

m5n.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

m5n.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3055m	-Xmx3055m
`mapreduce.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.map.memory.mb`	3819	3819
`mapreduce.reduce.memory.mb`	7638	7638
`yarn.app.mapreduce.am.resource.mb`	7638	7638
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30520
`yarn.nodemanager.resource.memory-mb`	366592	30520

m5zn 인스턴스

m5zn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2304m	-Xmx2304m
mapreduce.java.opts	-Xmx4608m	-Xmx4608m
mapreduce.map.memory.mb	2,880	2880
mapreduce.reduce.memory.mb	5760	5760
yarn.app.mapreduce.am.resource.mb	5760	5760
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11520	5760
yarn.nodemanager.resource.memory-mb	11520	5760

m5zn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m5zn.3xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2577m	-Xmx2577m
mapreduce.java.opts	-Xmx5154m	-Xmx5154m
mapreduce.map.memory.mb	3221	3221
mapreduce.reduce.memory.mb	6442	6442
yarn.app.mapreduce.am.resource.mb	6442	6442
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	38656	19328
yarn.nodemanager.resource.memory-mb	38656	19328

m5zn.6xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2850m	-Xmx2850m
mapreduce.java.opts	-Xmx5700m	-Xmx5700m
mapreduce.map.memory.mb	3563	3563
mapreduce.reduce.memory.mb	7126	7126
yarn.app.mapreduce.am.resource.mb	7126	7126
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	85504	28496
yarn.nodemanager.resource.memory-mb	85504	28496

m5zn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6a 인스턴스

m6a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2986m	-Xmx2986m
`mapreduce.java.opts`	-Xmx5972m	-Xmx5972m
`mapreduce.map.memory.mb`	3733	3733
`mapreduce.reduce.memory.mb`	7466	7466
`yarn.app.mapreduce.am.resource.mb`	7466	7466
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	179200	29880
`yarn.nodemanager.resource.memory-mb`	179200	29880

m6a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6a.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3072m	-Xmx3072m
`mapreduce.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.map.memory.mb`	3840	3840
`mapreduce.reduce.memory.mb`	7680	7680
`yarn.app.mapreduce.am.resource.mb`	7680	7680
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	491520	30720
`yarn.nodemanager.resource.memory-mb`	491520	30720

m6a.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3089m	-Xmx3089m
`mapreduce.java.opts`	-Xmx6178m	-Xmx6178m
`mapreduce.map.memory.mb`	3861	3861
`mapreduce.reduce.memory.mb`	7722	7722
`yarn.app.mapreduce.am.resource.mb`	7722	7722

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

m6g 인스턴스

m6g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

m6g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

m6gd 인스턴스

m6gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

m6gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m6gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6i 인스턴스

m6i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6i.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

m6i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	181248	30208
`yarn.nodemanager.resource.memory-mb`	181248	30208

m6i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

m6i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3055m	-Xmx3055m
`mapreduce.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.map.memory.mb`	3819	3819
`mapreduce.reduce.memory.mb`	7638	7638
`yarn.app.mapreduce.am.resource.mb`	7638	7638
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30520
`yarn.nodemanager.resource.memory-mb`	366592	30520

m6i.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m6id 인스턴스

m6id.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6id.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6id.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6id.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6id.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6id.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6id.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6id.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m6idn 인스턴스

m6idn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6idn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m6idn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6idn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6idn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m6idn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6idn.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3055m	-Xmx3055m
`mapreduce.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.map.memory.mb`	3819	3819
`mapreduce.reduce.memory.mb`	7638	7638
`yarn.app.mapreduce.am.resource.mb`	7638	7638
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30520
`yarn.nodemanager.resource.memory-mb`	366592	30520

m6idn.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3072m	-Xmx3072m
`mapreduce.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.map.memory.mb`	3840	3840
`mapreduce.reduce.memory.mb`	7680	7680
`yarn.app.mapreduce.am.resource.mb`	7680	7680

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m6in 인스턴스

m6in.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m6in.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

m6in.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m6in.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m6in.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	181248	30208
`yarn.nodemanager.resource.memory-mb`	181248	30208

m6in.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m6in.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m6in.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m7a 인스턴스

m7a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

m7a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3055m	-Xmx3055m
`mapreduce.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.map.memory.mb`	3819	3819
`mapreduce.reduce.memory.mb`	7638	7638
`yarn.app.mapreduce.am.resource.mb`	7638	7638

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7a.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3072m	-Xmx3072m
mapreduce.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.map.memory.mb	3840	3840
mapreduce.reduce.memory.mb	7680	7680
yarn.app.mapreduce.am.resource.mb	7680	7680
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

m7a.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3089m	-Xmx3089m
`mapreduce.java.opts`	-Xmx6178m	-Xmx6178m
`mapreduce.map.memory.mb`	3861	3861
`mapreduce.reduce.memory.mb`	7722	7722
`yarn.app.mapreduce.am.resource.mb`	7722	7722
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	741376	30952
`yarn.nodemanager.resource.memory-mb`	741376	30952

m7g 인스턴스

m7g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

m7g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.java.opts`	-Xmx5836m	-Xmx5836m
`mapreduce.map.memory.mb`	3648	3648
`mapreduce.reduce.memory.mb`	7296	7296
`yarn.app.mapreduce.am.resource.mb`	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

m7gd 인스턴스

m7gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

m7gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.java.opts`	-Xmx5836m	-Xmx5836m
`mapreduce.map.memory.mb`	3648	3648
`mapreduce.reduce.memory.mb`	7296	7296
`yarn.app.mapreduce.am.resource.mb`	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2986m	-Xmx2986m
mapreduce.java.opts	-Xmx5972m	-Xmx5972m
mapreduce.map.memory.mb	3733	3733
mapreduce.reduce.memory.mb	7466	7466
yarn.app.mapreduce.am.resource.mb	7466	7466
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	179200	29880
yarn.nodemanager.resource.memory-mb	179200	29880

m7gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m7i 인스턴스

m7i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m7i.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

m7i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.java.opts`	-Xmx5836m	-Xmx5836m
`mapreduce.map.memory.mb`	3648	3648
`mapreduce.reduce.memory.mb`	7296	7296
`yarn.app.mapreduce.am.resource.mb`	7296	7296

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m7i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m7i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

m7i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3055m	-Xmx3055m
`mapreduce.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.map.memory.mb`	3819	3819
`mapreduce.reduce.memory.mb`	7638	7638
`yarn.app.mapreduce.am.resource.mb`	7638	7638

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m7i.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3089m	-Xmx3089m
mapreduce.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.map.memory.mb	3861	3861
mapreduce.reduce.memory.mb	7722	7722
yarn.app.mapreduce.am.resource.mb	7722	7722
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

m7i-flex 인스턴스

m7i-flex.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	11712	5856
`yarn.nodemanager.resource.memory-mb`	11712	5856

m7i-flex.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

m7i-flex.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2714m	-Xmx2714m
mapreduce.java.opts	-Xmx5428m	-Xmx5428m
mapreduce.map.memory.mb	3392	3392
mapreduce.reduce.memory.mb	6784	6784
yarn.app.mapreduce.am.resource.mb	6784	6784
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m7i-flex.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2918m	-Xmx2918m
`mapreduce.java.opts`	-Xmx5836m	-Xmx5836m
`mapreduce.map.memory.mb`	3648	3648
`mapreduce.reduce.memory.mb`	7296	7296
`yarn.app.mapreduce.am.resource.mb`	7296	7296
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	116736	29184
`yarn.nodemanager.resource.memory-mb`	116736	29184

m7i-flex.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2986m	-Xmx2986m
`mapreduce.java.opts`	-Xmx5972m	-Xmx5972m
`mapreduce.map.memory.mb`	3733	3733
`mapreduce.reduce.memory.mb`	7466	7466
`yarn.app.mapreduce.am.resource.mb`	7466	7466

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	179200	29880
`yarn.nodemanager.resource.memory-mb`	179200	29880

m7i-flex.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

m8g 인스턴스

m8g.xLarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m8g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

m8g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

m8g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m8g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	181248	30208
yarn.nodemanager.resource.memory-mb	181248	30208

m8g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m8g.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3055m	-Xmx3055m
`mapreduce.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.map.memory.mb`	3819	3819
`mapreduce.reduce.memory.mb`	7638	7638
`yarn.app.mapreduce.am.resource.mb`	7638	7638
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30520
`yarn.nodemanager.resource.memory-mb`	366592	30520

m8g.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3089m	-Xmx3089m
`mapreduce.java.opts`	-Xmx6178m	-Xmx6178m
`mapreduce.map.memory.mb`	3861	3861
`mapreduce.reduce.memory.mb`	7722	7722
`yarn.app.mapreduce.am.resource.mb`	7722	7722

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30952
yarn.nodemanager.resource.memory-mb	741376	30952

m8gd 인스턴스

m8gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2342m	-Xmx2342m
mapreduce.java.opts	-Xmx4684m	-Xmx4684m
mapreduce.map.memory.mb	2928	2928
mapreduce.reduce.memory.mb	5856	5856
yarn.app.mapreduce.am.resource.mb	5856	5856
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	11712	5856
yarn.nodemanager.resource.memory-mb	11712	5856

m8gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

m8gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

m8gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

m8gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2986m	-Xmx2986m
`mapreduce.java.opts`	-Xmx5972m	-Xmx5972m
`mapreduce.map.memory.mb`	3733	3733
`mapreduce.reduce.memory.mb`	7466	7466
`yarn.app.mapreduce.am.resource.mb`	7466	7466
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	179200	29880
`yarn.nodemanager.resource.memory-mb`	179200	29880

m8gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3021m	-Xmx3021m
`mapreduce.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.map.memory.mb`	3776	3776
`mapreduce.reduce.memory.mb`	7552	7552
`yarn.app.mapreduce.am.resource.mb`	7552	7552

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

m8gd.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3055m	-Xmx3055m
mapreduce.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.map.memory.mb	3819	3819
mapreduce.reduce.memory.mb	7638	7638
yarn.app.mapreduce.am.resource.mb	7638	7638
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30520
yarn.nodemanager.resource.memory-mb	366592	30520

m8gd.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx3089m	-Xmx3089m
`mapreduce.java.opts`	-Xmx6178m	-Xmx6178m
`mapreduce.map.memory.mb`	3861	3861
`mapreduce.reduce.memory.mb`	7722	7722
`yarn.app.mapreduce.am.resource.mb`	7722	7722
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	741376	30952
`yarn.nodemanager.resource.memory-mb`	741376	30952

p2 인스턴스

p2.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.java.opts`	-Xmx21708m	-Xmx21708m
`mapreduce.map.memory.mb`	13568	13568
`mapreduce.reduce.memory.mb`	27136	27136
`yarn.app.mapreduce.am.resource.mb`	27136	27136

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p2.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

p2.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx9267m	-Xmx9267m
`mapreduce.java.opts`	-Xmx18534m	-Xmx18534m
`mapreduce.map.memory.mb`	11584	11584
`mapreduce.reduce.memory.mb`	23168	23168
`yarn.app.mapreduce.am.resource.mb`	23168	23168
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	741376	23168
`yarn.nodemanager.resource.memory-mb`	741376	23168

p3 인스턴스

p3.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5427m	-Xmx5427m
`mapreduce.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.map.memory.mb`	6784	6784
`mapreduce.reduce.memory.mb`	13568	13568
`yarn.app.mapreduce.am.resource.mb`	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

p3.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

p3.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

p4d 인스턴스

p4d.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx9302m	-Xmx9302m
mapreduce.java.opts	-Xmx18604m	-Xmx18604m
mapreduce.map.memory.mb	11627	11627
mapreduce.reduce.memory.mb	23254	23254
yarn.app.mapreduce.am.resource.mb	23254	23254

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1116160	23222
yarn.nodemanager.resource.memory-mb	1116160	23222

p5 인스턴스

p5.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx8294m	-Xmx8294m
mapreduce.java.opts	-Xmx16588m	-Xmx16588m
mapreduce.map.memory.mb	10368	10368
mapreduce.reduce.memory.mb	20736	20736
yarn.app.mapreduce.am.resource.mb	20736	20736
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	20736
yarn.nodemanager.resource.memory-mb	1990656	20736

r3 인스턴스

r3.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2342m	-Xmx2342m
`mapreduce.java.opts`	-Xmx4684m	-Xmx4684m
`mapreduce.map.memory.mb`	2928	2928
`mapreduce.reduce.memory.mb`	5856	5856
`yarn.app.mapreduce.am.resource.mb`	5856	5856
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

r3.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx2714m	-Xmx2714m
`mapreduce.java.opts`	-Xmx5428m	-Xmx5428m
`mapreduce.map.memory.mb`	3392	3392
`mapreduce.reduce.memory.mb`	6784	6784
`yarn.app.mapreduce.am.resource.mb`	6784	6784

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r3.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx2918m	-Xmx2918m
mapreduce.java.opts	-Xmx5836m	-Xmx5836m
mapreduce.map.memory.mb	3648	3648
mapreduce.reduce.memory.mb	7296	7296
yarn.app.mapreduce.am.resource.mb	7296	7296
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r3.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx3021m	-Xmx3021m
mapreduce.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.map.memory.mb	3776	3776
mapreduce.reduce.memory.mb	7552	7552
yarn.app.mapreduce.am.resource.mb	7552	7552
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r4 인스턴스

r4.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r4.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r4.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r4.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r4.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5 인스턴스

r5.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

r5a 인스턴스

r5a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.java.opts`	-Xmx12288m	-Xmx12288m
`mapreduce.map.memory.mb`	7680	7680
`mapreduce.reduce.memory.mb`	15360	15360
`yarn.app.mapreduce.am.resource.mb`	15360	15360
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	122880	30720
`yarn.nodemanager.resource.memory-mb`	122880	30720

r5a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6349m	-Xmx6349m
`mapreduce.java.opts`	-Xmx12698m	-Xmx12698m
`mapreduce.map.memory.mb`	7936	7936
`mapreduce.reduce.memory.mb`	15872	15872
`yarn.app.mapreduce.am.resource.mb`	15872	15872

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	253952	31744
`yarn.nodemanager.resource.memory-mb`	253952	31744

r5a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6417m	-Xmx6417m
`mapreduce.java.opts`	-Xmx12834m	-Xmx12834m
`mapreduce.map.memory.mb`	8021	8021
`mapreduce.reduce.memory.mb`	16042	16042
`yarn.app.mapreduce.am.resource.mb`	16042	16042
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	385024	32100
`yarn.nodemanager.resource.memory-mb`	385024	32100

r5a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

r5ad 인스턴스

r5ad.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5ad.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5427m	-Xmx5427m
`mapreduce.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.map.memory.mb`	6784	6784
`mapreduce.reduce.memory.mb`	13568	13568
`yarn.app.mapreduce.am.resource.mb`	13568	13568
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

r5ad.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5ad.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5ad.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.java.opts`	-Xmx12220m	-Xmx12220m
`mapreduce.map.memory.mb`	7637	7637
`mapreduce.reduce.memory.mb`	15274	15274
`yarn.app.mapreduce.am.resource.mb`	15274	15274
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30564
`yarn.nodemanager.resource.memory-mb`	366592	30564

r5ad.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6246m	-Xmx6246m
`mapreduce.java.opts`	-Xmx12492m	-Xmx12492m
`mapreduce.map.memory.mb`	7808	7808
`mapreduce.reduce.memory.mb`	15616	15616
`yarn.app.mapreduce.am.resource.mb`	15616	15616

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	499712	31232
yarn.nodemanager.resource.memory-mb	499712	31232

r5ad.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r5b 인스턴스

r5b.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5b.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5b.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5b.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5b.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5b.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5b.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r5d 인스턴스

r5d.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

r5d.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

r5d.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	122880	30720
yarn.nodemanager.resource.memory-mb	122880	30720

r5d.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6349m	-Xmx6349m
mapreduce.java.opts	-Xmx12698m	-Xmx12698m
mapreduce.map.memory.mb	7936	7936
mapreduce.reduce.memory.mb	15872	15872
yarn.app.mapreduce.am.resource.mb	15872	15872

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	253952	31744
yarn.nodemanager.resource.memory-mb	253952	31744

r5d.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

r5d.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6451m	-Xmx6451m
mapreduce.java.opts	-Xmx12902m	-Xmx12902m
mapreduce.map.memory.mb	8064	8064
mapreduce.reduce.memory.mb	16128	16128
yarn.app.mapreduce.am.resource.mb	16128	16128
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	516096	32256
yarn.nodemanager.resource.memory-mb	516096	32256

r5d.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6486m	-Xmx6486m
mapreduce.java.opts	-Xmx12972m	-Xmx12972m
mapreduce.map.memory.mb	8107	8107
mapreduce.reduce.memory.mb	16214	16214
yarn.app.mapreduce.am.resource.mb	16214	16214

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	778240	32396
yarn.nodemanager.resource.memory-mb	778240	32396

r5dn 인스턴스

r5dn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5dn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5427m	-Xmx5427m
`mapreduce.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.map.memory.mb`	6784	6784
`mapreduce.reduce.memory.mb`	13568	13568
`yarn.app.mapreduce.am.resource.mb`	13568	13568
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

r5dn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5dn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r5dn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5dn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5dn.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r5n 인스턴스

r5n.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r5n.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r5n.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r5n.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.java.opts`	-Xmx12084m	-Xmx12084m
`mapreduce.map.memory.mb`	7552	7552
`mapreduce.reduce.memory.mb`	15104	15104
`yarn.app.mapreduce.am.resource.mb`	15104	15104
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

r5n.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.java.opts`	-Xmx12220m	-Xmx12220m
`mapreduce.map.memory.mb`	7637	7637
`mapreduce.reduce.memory.mb`	15274	15274
`yarn.app.mapreduce.am.resource.mb`	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r5n.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r5n.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6a 인스턴스

r6a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6a.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r6a.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

r6g 인스턴스

r6g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	116736	29184
`yarn.nodemanager.resource.memory-mb`	116736	29184

r6g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.java.opts`	-Xmx12084m	-Xmx12084m
`mapreduce.map.memory.mb`	7552	7552
`mapreduce.reduce.memory.mb`	15104	15104
`yarn.app.mapreduce.am.resource.mb`	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6gd 인스턴스

r6gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

r6gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.java.opts`	-Xmx12220m	-Xmx12220m
`mapreduce.map.memory.mb`	7637	7637
`mapreduce.reduce.memory.mb`	15274	15274
`yarn.app.mapreduce.am.resource.mb`	15274	15274
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30564
`yarn.nodemanager.resource.memory-mb`	366592	30564

r6gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6i 인스턴스

r6i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.java.opts`	-Xmx12288m	-Xmx12288m
`mapreduce.map.memory.mb`	7680	7680
`mapreduce.reduce.memory.mb`	15360	15360
`yarn.app.mapreduce.am.resource.mb`	15360	15360
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	491520	30720
`yarn.nodemanager.resource.memory-mb`	491520	30720

r6i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6178m	-Xmx6178m
`mapreduce.java.opts`	-Xmx12356m	-Xmx12356m
`mapreduce.map.memory.mb`	7723	7723
`mapreduce.reduce.memory.mb`	15446	15446
`yarn.app.mapreduce.am.resource.mb`	15446	15446

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6i.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

r6id 인스턴스

r6id.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6id.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6id.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6id.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6id.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6id.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6id.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6id.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r6idn 인스턴스

r6idn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6idn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5427m	-Xmx5427m
`mapreduce.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.map.memory.mb`	6784	6784
`mapreduce.reduce.memory.mb`	13568	13568
`yarn.app.mapreduce.am.resource.mb`	13568	13568
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

r6idn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r6idn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6idn.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.java.opts`	-Xmx12220m	-Xmx12220m
`mapreduce.map.memory.mb`	7637	7637
`mapreduce.reduce.memory.mb`	15274	15274
`yarn.app.mapreduce.am.resource.mb`	15274	15274
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30564
`yarn.nodemanager.resource.memory-mb`	366592	30564

r6idn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.java.opts`	-Xmx12288m	-Xmx12288m
`mapreduce.map.memory.mb`	7680	7680
`mapreduce.reduce.memory.mb`	15360	15360
`yarn.app.mapreduce.am.resource.mb`	15360	15360

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r6idn.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6idn.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6195m	-Xmx6195m
`mapreduce.java.opts`	-Xmx12390m	-Xmx12390m
`mapreduce.map.memory.mb`	7744	7744
`mapreduce.reduce.memory.mb`	15488	15488
`yarn.app.mapreduce.am.resource.mb`	15488	15488
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	991232	30976
`yarn.nodemanager.resource.memory-mb`	991232	30976

r6in 인스턴스

r6in.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx4685m	-Xmx4685m
`mapreduce.java.opts`	-Xmx9370m	-Xmx9370m
`mapreduce.map.memory.mb`	5856	5856
`mapreduce.reduce.memory.mb`	11712	11712
`yarn.app.mapreduce.am.resource.mb`	11712	11712

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r6in.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r6in.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx5837m	-Xmx5837m
`mapreduce.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.map.memory.mb`	7296	7296
`mapreduce.reduce.memory.mb`	14592	14592
`yarn.app.mapreduce.am.resource.mb`	14592	14592
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	116736	29184
`yarn.nodemanager.resource.memory-mb`	116736	29184

r6in.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6042m	-Xmx6042m
`mapreduce.java.opts`	-Xmx12084m	-Xmx12084m
`mapreduce.map.memory.mb`	7552	7552
`mapreduce.reduce.memory.mb`	15104	15104
`yarn.app.mapreduce.am.resource.mb`	15104	15104

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r6in.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r6in.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.java.opts`	-Xmx12288m	-Xmx12288m
`mapreduce.map.memory.mb`	7680	7680
`mapreduce.reduce.memory.mb`	15360	15360
`yarn.app.mapreduce.am.resource.mb`	15360	15360
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	491520	30720
`yarn.nodemanager.resource.memory-mb`	491520	30720

r6in.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6178m	-Xmx6178m
`mapreduce.java.opts`	-Xmx12356m	-Xmx12356m
`mapreduce.map.memory.mb`	7723	7723
`mapreduce.reduce.memory.mb`	15446	15446
`yarn.app.mapreduce.am.resource.mb`	15446	15446

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r6in.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r7a 인스턴스

r7a.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7a.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7a.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7a.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7a.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7a.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7a.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7a.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6195m	-Xmx6195m
mapreduce.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.map.memory.mb	7744	7744
mapreduce.reduce.memory.mb	15488	15488
yarn.app.mapreduce.am.resource.mb	15488	15488

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

r7a.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

r7g 인스턴스

r7g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7gd 인스턴스

r7gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7i 인스턴스

r7i.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r7i.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7i.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7i.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7i.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7i.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7i.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r7i.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

구성 옵션	기본값	HBase가 설치된 경우
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	1490944	31124
`yarn.nodemanager.resource.memory-mb`	1490944	31124

r7iz 인스턴스

r7iz.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx4685m	-Xmx4685m
`mapreduce.java.opts`	-Xmx9370m	-Xmx9370m
`mapreduce.map.memory.mb`	5856	5856
`mapreduce.reduce.memory.mb`	11712	11712
`yarn.app.mapreduce.am.resource.mb`	11712	11712
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	23424	11712
`yarn.nodemanager.resource.memory-mb`	23424	11712

r7iz.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r7iz.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r7iz.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r7iz.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r7iz.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r7iz.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6029m	-Xmx6029m
mapreduce.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.map.memory.mb	7536	7536
mapreduce.reduce.memory.mb	15072	15072
yarn.app.mapreduce.am.resource.mb	15072	15072
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

r8g 인스턴스

r8g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r8g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r8g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r8g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r8g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6110m	-Xmx6110m
mapreduce.java.opts	-Xmx12220m	-Xmx12220m
mapreduce.map.memory.mb	7637	7637
mapreduce.reduce.memory.mb	15274	15274
yarn.app.mapreduce.am.resource.mb	15274	15274

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	366592	30564
yarn.nodemanager.resource.memory-mb	366592	30564

r8g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6144m	-Xmx6144m
mapreduce.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.map.memory.mb	7680	7680
mapreduce.reduce.memory.mb	15360	15360
yarn.app.mapreduce.am.resource.mb	15360	15360
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r8g.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r8g.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

r8gd 인스턴스

r8gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4685m	-Xmx4685m
mapreduce.java.opts	-Xmx9370m	-Xmx9370m
mapreduce.map.memory.mb	5856	5856
mapreduce.reduce.memory.mb	11712	11712
yarn.app.mapreduce.am.resource.mb	11712	11712
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	23424	11712
yarn.nodemanager.resource.memory-mb	23424	11712

r8gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5427m	-Xmx5427m
mapreduce.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.map.memory.mb	6784	6784
mapreduce.reduce.memory.mb	13568	13568
yarn.app.mapreduce.am.resource.mb	13568	13568
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

r8gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5837m	-Xmx5837m
mapreduce.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.map.memory.mb	7296	7296
mapreduce.reduce.memory.mb	14592	14592
yarn.app.mapreduce.am.resource.mb	14592	14592

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

r8gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6042m	-Xmx6042m
mapreduce.java.opts	-Xmx12084m	-Xmx12084m
mapreduce.map.memory.mb	7552	7552
mapreduce.reduce.memory.mb	15104	15104
yarn.app.mapreduce.am.resource.mb	15104	15104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

r8gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6110m	-Xmx6110m
`mapreduce.java.opts`	-Xmx12220m	-Xmx12220m
`mapreduce.map.memory.mb`	7637	7637
`mapreduce.reduce.memory.mb`	15274	15274
`yarn.app.mapreduce.am.resource.mb`	15274	15274
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	366592	30564
`yarn.nodemanager.resource.memory-mb`	366592	30564

r8gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx6144m	-Xmx6144m
`mapreduce.java.opts`	-Xmx12288m	-Xmx12288m
`mapreduce.map.memory.mb`	7680	7680
`mapreduce.reduce.memory.mb`	15360	15360
`yarn.app.mapreduce.am.resource.mb`	15360	15360

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

r8gd.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6178m	-Xmx6178m
mapreduce.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.map.memory.mb	7723	7723
mapreduce.reduce.memory.mb	15446	15446
yarn.app.mapreduce.am.resource.mb	15446	15446
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30860
yarn.nodemanager.resource.memory-mb	741376	30860

r8gd.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6212m	-Xmx6212m
mapreduce.java.opts	-Xmx12424m	-Xmx12424m
mapreduce.map.memory.mb	7765	7765
mapreduce.reduce.memory.mb	15530	15530
yarn.app.mapreduce.am.resource.mb	15530	15530
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31124
yarn.nodemanager.resource.memory-mb	1490944	31124

x1 인스턴스

x1.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12058m	-Xmx12058m
mapreduce.java.opts	-Xmx24116m	-Xmx24116m
mapreduce.map.memory.mb	15072	15072
mapreduce.reduce.memory.mb	30144	30144
yarn.app.mapreduce.am.resource.mb	30144	30144

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	30144
yarn.nodemanager.resource.memory-mb	964608	30144

x1.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12109m	-Xmx12109m
mapreduce.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.map.memory.mb	15136	15136
mapreduce.reduce.memory.mb	30272	30272
yarn.app.mapreduce.am.resource.mb	30272	30272
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937408	30272
yarn.nodemanager.resource.memory-mb	1937408	30272

x1e 인스턴스

x1e.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx22682m	-Xmx22682m
mapreduce.java.opts	-Xmx45364m	-Xmx45364m
mapreduce.map.memory.mb	28352	28352
mapreduce.reduce.memory.mb	56704	56704
yarn.app.mapreduce.am.resource.mb	56704	56704
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	113408	0
yarn.nodemanager.resource.memory-mb	113408	0

x1e.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx23501m	-Xmx23501m
mapreduce.java.opts	-Xmx47002m	-Xmx47002m
mapreduce.map.memory.mb	29376	29376
mapreduce.reduce.memory.mb	58752	58752
yarn.app.mapreduce.am.resource.mb	58752	58752

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	235008	0
yarn.nodemanager.resource.memory-mb	235008	0

x1e.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx23910m	-Xmx23910m
mapreduce.java.opts	-Xmx47820m	-Xmx47820m
mapreduce.map.memory.mb	29888	29888
mapreduce.reduce.memory.mb	59776	59776
yarn.app.mapreduce.am.resource.mb	59776	59776
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	478208	0
yarn.nodemanager.resource.memory-mb	478208	0

x1e.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx24115m	-Xmx24115m
mapreduce.java.opts	-Xmx48230m	-Xmx48230m
mapreduce.map.memory.mb	30144	30144
mapreduce.reduce.memory.mb	60288	60288
yarn.app.mapreduce.am.resource.mb	60288	60288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	964608	0
yarn.nodemanager.resource.memory-mb	964608	0

x1e.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx24218m	-Xmx24218m
mapreduce.java.opts	-Xmx48436m	-Xmx48436m
mapreduce.map.memory.mb	30272	30272
mapreduce.reduce.memory.mb	60544	60544
yarn.app.mapreduce.am.resource.mb	60544	60544

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1937408	0
yarn.nodemanager.resource.memory-mb	1937408	0

x1e.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx24269m	-Xmx24269m
mapreduce.java.opts	-Xmx48538m	-Xmx48538m
mapreduce.map.memory.mb	30336	30336
mapreduce.reduce.memory.mb	60672	60672
yarn.app.mapreduce.am.resource.mb	60672	60672
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3883008	0
yarn.nodemanager.resource.memory-mb	3883008	0

x2gd 인스턴스

x2gd.xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx10854m	-Xmx10854m
`mapreduce.java.opts`	-Xmx21708m	-Xmx21708m
`mapreduce.map.memory.mb`	13568	13568
`mapreduce.reduce.memory.mb`	27136	27136
`yarn.app.mapreduce.am.resource.mb`	27136	27136
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	54272	27136
`yarn.nodemanager.resource.memory-mb`	54272	27136

x2gd.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx11674m	-Xmx11674m
`mapreduce.java.opts`	-Xmx23348m	-Xmx23348m
`mapreduce.map.memory.mb`	14592	14592
`mapreduce.reduce.memory.mb`	29184	29184
`yarn.app.mapreduce.am.resource.mb`	29184	29184

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

x2gd.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12083m	-Xmx12083m
mapreduce.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.map.memory.mb	15104	15104
mapreduce.reduce.memory.mb	30208	30208
yarn.app.mapreduce.am.resource.mb	30208	30208
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	30208
yarn.nodemanager.resource.memory-mb	241664	30208

x2gd.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12288m	-Xmx12288m
mapreduce.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.map.memory.mb	15360	15360
mapreduce.reduce.memory.mb	30720	30720
yarn.app.mapreduce.am.resource.mb	30720	30720
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

x2gd.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

x2gd.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12390m	-Xmx12390m
mapreduce.java.opts	-Xmx24780m	-Xmx24780m
mapreduce.map.memory.mb	15488	15488
mapreduce.reduce.memory.mb	30976	30976
yarn.app.mapreduce.am.resource.mb	30976	30976
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	991232	30976
yarn.nodemanager.resource.memory-mb	991232	30976

x2idn 인스턴스

x2idn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx12390m	-Xmx12390m
`mapreduce.java.opts`	-Xmx24780m	-Xmx24780m
`mapreduce.map.memory.mb`	15488	15488
`mapreduce.reduce.memory.mb`	30976	30976
`yarn.app.mapreduce.am.resource.mb`	30976	30976
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	991232	30976
`yarn.nodemanager.resource.memory-mb`	991232	30976

x2idn.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx12425m	-Xmx12425m
`mapreduce.java.opts`	-Xmx24850m	-Xmx24850m
`mapreduce.map.memory.mb`	15531	15531
`mapreduce.reduce.memory.mb`	31062	31062
`yarn.app.mapreduce.am.resource.mb`	31062	31062

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

x2idn.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12442m	-Xmx12442m
mapreduce.java.opts	-Xmx24884m	-Xmx24884m
mapreduce.map.memory.mb	15552	15552
mapreduce.reduce.memory.mb	31104	31104
yarn.app.mapreduce.am.resource.mb	31104	31104
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	31104
yarn.nodemanager.resource.memory-mb	1990656	31104

x2iedn 인스턴스

x2iedn.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx23347m	-Xmx23347m
mapreduce.java.opts	-Xmx46694m	-Xmx46694m
mapreduce.map.memory.mb	29184	29184
mapreduce.reduce.memory.mb	58368	58368
yarn.app.mapreduce.am.resource.mb	58368	58368
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	0
yarn.nodemanager.resource.memory-mb	116736	0

x2iedn.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx24166m	-Xmx24166m
mapreduce.java.opts	-Xmx48332m	-Xmx48332m
mapreduce.map.memory.mb	30208	30208
mapreduce.reduce.memory.mb	60416	60416
yarn.app.mapreduce.am.resource.mb	60416	60416

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	241664	0
yarn.nodemanager.resource.memory-mb	241664	0

x2iedn.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx24576m	-Xmx24576m
mapreduce.java.opts	-Xmx49152m	-Xmx49152m
mapreduce.map.memory.mb	30720	30720
mapreduce.reduce.memory.mb	61440	61440
yarn.app.mapreduce.am.resource.mb	61440	61440
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	0
yarn.nodemanager.resource.memory-mb	491520	0

x2iedn.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx24781m	-Xmx24781m
`mapreduce.java.opts`	-Xmx49562m	-Xmx49562m
`mapreduce.map.memory.mb`	30976	30976
`mapreduce.reduce.memory.mb`	61952	61952
`yarn.app.mapreduce.am.resource.mb`	61952	61952
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	991232	0
`yarn.nodemanager.resource.memory-mb`	991232	0

x2iedn.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx24883m	-Xmx24883m
`mapreduce.java.opts`	-Xmx49766m	-Xmx49766m
`mapreduce.map.memory.mb`	31104	31104
`mapreduce.reduce.memory.mb`	62208	62208
`yarn.app.mapreduce.am.resource.mb`	62208	62208

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1990656	0
yarn.nodemanager.resource.memory-mb	1990656	0

x2iedn.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx24918m	-Xmx24918m
mapreduce.java.opts	-Xmx49836m	-Xmx49836m
mapreduce.map.memory.mb	31147	31147
mapreduce.reduce.memory.mb	62294	62294
yarn.app.mapreduce.am.resource.mb	62294	62294
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	-32
yarn.nodemanager.resource.memory-mb	2990080	-32

x2iedn.32xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx24934m	-Xmx24934m
mapreduce.java.opts	-Xmx49868m	-Xmx49868m
mapreduce.map.memory.mb	31168	31168
mapreduce.reduce.memory.mb	62336	62336
yarn.app.mapreduce.am.resource.mb	62336	62336
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	3989504	0
yarn.nodemanager.resource.memory-mb	3989504	0

x8g 인스턴스

x8g.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx10854m	-Xmx10854m
mapreduce.java.opts	-Xmx21708m	-Xmx21708m
mapreduce.map.memory.mb	13568	13568
mapreduce.reduce.memory.mb	27136	27136
yarn.app.mapreduce.am.resource.mb	27136	27136

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	54272	27136
yarn.nodemanager.resource.memory-mb	54272	27136

x8g.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx11674m	-Xmx11674m
mapreduce.java.opts	-Xmx23348m	-Xmx23348m
mapreduce.map.memory.mb	14592	14592
mapreduce.reduce.memory.mb	29184	29184
yarn.app.mapreduce.am.resource.mb	29184	29184
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	116736	29184
yarn.nodemanager.resource.memory-mb	116736	29184

x8g.4xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx12083m	-Xmx12083m
`mapreduce.java.opts`	-Xmx24166m	-Xmx24166m
`mapreduce.map.memory.mb`	15104	15104
`mapreduce.reduce.memory.mb`	30208	30208
`yarn.app.mapreduce.am.resource.mb`	30208	30208
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	241664	30208
`yarn.nodemanager.resource.memory-mb`	241664	30208

x8g.8xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx12288m	-Xmx12288m
`mapreduce.java.opts`	-Xmx24576m	-Xmx24576m
`mapreduce.map.memory.mb`	15360	15360
`mapreduce.reduce.memory.mb`	30720	30720
`yarn.app.mapreduce.am.resource.mb`	30720	30720

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	491520	30720
yarn.nodemanager.resource.memory-mb	491520	30720

x8g.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12356m	-Xmx12356m
mapreduce.java.opts	-Xmx24712m	-Xmx24712m
mapreduce.map.memory.mb	15445	15445
mapreduce.reduce.memory.mb	30890	30890
yarn.app.mapreduce.am.resource.mb	30890	30890
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	741376	30906
yarn.nodemanager.resource.memory-mb	741376	30906

x8g.16xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx12390m	-Xmx12390m
`mapreduce.java.opts`	-Xmx24780m	-Xmx24780m
`mapreduce.map.memory.mb`	15488	15488
`mapreduce.reduce.memory.mb`	30976	30976
`yarn.app.mapreduce.am.resource.mb`	30976	30976
`yarn.scheduler.minimum-allocation-mb`	1	1
`yarn.scheduler.maximum-allocation-mb`	991232	30976
`yarn.nodemanager.resource.memory-mb`	991232	30976

x8g.24xlarge

구성 옵션	기본값	HBase가 설치된 경우
`mapreduce.map.java.opts`	-Xmx12425m	-Xmx12425m
`mapreduce.java.opts`	-Xmx24850m	-Xmx24850m
`mapreduce.map.memory.mb`	15531	15531
`mapreduce.reduce.memory.mb`	31062	31062
`yarn.app.mapreduce.am.resource.mb`	31062	31062

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	1490944	31030
yarn.nodemanager.resource.memory-mb	1490944	31030

x8g.48xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx12458m	-Xmx12458m
mapreduce.java.opts	-Xmx24916m	-Xmx24916m
mapreduce.map.memory.mb	15573	15573
mapreduce.reduce.memory.mb	31146	31146
yarn.app.mapreduce.am.resource.mb	31146	31146
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	2990080	31210
yarn.nodemanager.resource.memory-mb	2990080	31210

z1d 인스턴스

z1d.xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx4915m	-Xmx4915m
mapreduce.java.opts	-Xmx9830m	-Xmx9830m
mapreduce.map.memory.mb	6144	6144
mapreduce.reduce.memory.mb	12288	12288
yarn.app.mapreduce.am.resource.mb	12288	12288
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	24576	12288
yarn.nodemanager.resource.memory-mb	24576	12288

z1d.2xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx5734m	-Xmx5734m
mapreduce.java.opts	-Xmx11468m	-Xmx11468m
mapreduce.map.memory.mb	7168	7168
mapreduce.reduce.memory.mb	14336	14336
yarn.app.mapreduce.am.resource.mb	14336	14336

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	57344	28672
yarn.nodemanager.resource.memory-mb	57344	28672

z1d.3xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6007m	-Xmx6007m
mapreduce.java.opts	-Xmx12014m	-Xmx12014m
mapreduce.map.memory.mb	7509	7509
mapreduce.reduce.memory.mb	15018	15018
yarn.app.mapreduce.am.resource.mb	15018	15018
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	90112	30040
yarn.nodemanager.resource.memory-mb	90112	30040

z1d.6xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6281m	-Xmx6281m
mapreduce.java.opts	-Xmx12562m	-Xmx12562m
mapreduce.map.memory.mb	7851	7851
mapreduce.reduce.memory.mb	15702	15702
yarn.app.mapreduce.am.resource.mb	15702	15702
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	188416	31396
yarn.nodemanager.resource.memory-mb	188416	31396

z1d.12xlarge

구성 옵션	기본값	HBase가 설치된 경우
mapreduce.map.java.opts	-Xmx6417m	-Xmx6417m
mapreduce.java.opts	-Xmx12834m	-Xmx12834m
mapreduce.map.memory.mb	8021	8021
mapreduce.reduce.memory.mb	16042	16042
yarn.app.mapreduce.am.resource.mb	16042	16042

구성 옵션	기본값	HBase가 설치된 경우
yarn.scheduler.minimum-allocation-mb	1	1
yarn.scheduler.maximum-allocation-mb	385024	32100
yarn.nodemanager.resource.memory-mb	385024	32100

Hadoop 대몬(daemon) 구성 설정

하둡 데몬 설정은 클러스터 노드에서 사용하는 EC2 인스턴스 유형에 따라 달라집니다. 다음 표에는 각 EC2 인스턴스 유형에 대한 기본 구성 설정이 나열되어 있습니다.

이 설정을 사용자 지정하려면 hadoop-env 구성 분류를 사용합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

인스턴스 유형

- c1 인스턴스
- c3 인스턴스
- c4 인스턴스
- c5 인스턴스
- c5a 인스턴스
- c5ad 인스턴스
- c5d 인스턴스
- c5n 인스턴스
- c6a 인스턴스
- c6g 인스턴스
- c6gd 인스턴스
- c6gn 인스턴스
- c6i 인스턴스
- c6id 인스턴스

- c6in 인스턴스
- c7a 인스턴스
- c7g 인스턴스
- c7gd 인스턴스
- c7gn 인스턴스
- c7i 인스턴스
- c7i-flex 인스턴스
- c8g 인스턴스
- c8gd 인스턴스
- d2 인스턴스
- d3 인스턴스
- d3en 인스턴스
- f2 인스턴스
- g3 인스턴스
- g3s 인스턴스
- g4dn 인스턴스
- g5 인스턴스
- g6 인스턴스
- g6e 인스턴스
- gr6 인스턴스
- h1 인스턴스
- i2 인스턴스
- i3 인스턴스
- i3en 인스턴스
- i4g 인스턴스
- i4i 인스턴스
- i7i 인스턴스
- i7ie 인스턴스
- i8g 인스턴스

- im4gn 인스턴스
- is4gen 인스턴스
- m1 인스턴스
- m2 인스턴스
- m3 인스턴스
- m4 인스턴스
- m5 인스턴스
- m5a 인스턴스
- m5ad 인스턴스
- m5d 인스턴스
- m5dn 인스턴스
- m5n 인스턴스
- m5zn 인스턴스
- m6a 인스턴스
- m6g 인스턴스
- m6gd 인스턴스
- m6i 인스턴스
- m6id 인스턴스
- m6idn 인스턴스
- m6in 인스턴스
- m7a 인스턴스
- m7g 인스턴스
- m7gd 인스턴스
- m7i 인스턴스
- m7i-flex 인스턴스
- m8g 인스턴스
- m8gd 인스턴스
- p2 인스턴스
- p3 인스턴스

- p4d 인스턴스
- p5 인스턴스
- r3 인스턴스
- r4 인스턴스
- r5 인스턴스
- r5a 인스턴스
- r5ad 인스턴스
- r5b 인스턴스
- r5d 인스턴스
- r5dn 인스턴스
- r5n 인스턴스
- r6a 인스턴스
- r6g 인스턴스
- r6gd 인스턴스
- r6i 인스턴스
- r6id 인스턴스
- r6idn 인스턴스
- r6in 인스턴스
- r7a 인스턴스
- r7g 인스턴스
- r7gd 인스턴스
- r7i 인스턴스
- r7iz 인스턴스
- r8g 인스턴스
- r8gd 인스턴스
- x1 인스턴스
- x1e 인스턴스
- x2gd 인스턴스
- x2idn 인스턴스

- x2iedn 인스턴스
- x8g 인스턴스
- z1d 인스턴스

c1 인스턴스

c1.medium

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	192
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	128
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

c1.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	768
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768

파라미터	값
HADOOP_DATANODE_HEAPSIZE	384

c3 인스턴스

c3.xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c3.2xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740

파라미터	값
HADOOP_DATANODE_HEAPSIZE	757

c3.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c3.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

c4 인스턴스

c4.large

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	1152
YARN_PROXYSERVER_HEAPSIZE	1152
YARN_NODEMANAGER_HEAPSIZE	1152
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1152
HADOOP_NAMENODE_HEAPSIZE	576
HADOOP_DATANODE_HEAPSIZE	384

c4.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c4.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

c4.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

c4.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

c5 인스턴스

c5.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5.9xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5.18xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

c5a 인스턴스

c5a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c5ad 인스턴스

c5ad.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c5ad.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c5ad.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c5ad.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c5ad.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c5ad.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c5ad.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c5d 인스턴스

c5d.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

c5d.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

c5d.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

c5d.9xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3563
YARN_PROXYSERVER_HEAPSIZE	3563
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3563
HADOOP_NAMENODE_HEAPSIZE	7577
HADOOP_DATANODE_HEAPSIZE	1925

c5d.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5d.18xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5038
YARN_PROXYSERVER_HEAPSIZE	5038
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5038
HADOOP_NAMENODE_HEAPSIZE	14950
HADOOP_DATANODE_HEAPSIZE	3399

c5d.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

c5n 인스턴스

c5n.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2304
YARN_PROXYSERVER_HEAPSIZE	2304
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2304
HADOOP_NAMENODE_HEAPSIZE	1280
HADOOP_DATANODE_HEAPSIZE	665

c5n.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2519
YARN_PROXYSERVER_HEAPSIZE	2519
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2519
HADOOP_NAMENODE_HEAPSIZE	2355
HADOOP_DATANODE_HEAPSIZE	880

c5n.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2949
YARN_PROXYSERVER_HEAPSIZE	2949
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2949
HADOOP_NAMENODE_HEAPSIZE	4505
HADOOP_DATANODE_HEAPSIZE	1310

c5n.9xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

c5n.18xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

c6a 인스턴스

c6a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6a.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c6a.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

c6g 인스턴스

c6g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6gd 인스턴스

c6gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6gn 인스턴스

c6gn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6gn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6gn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6gn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6gn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6gn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6i 인스턴스

c6i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6i.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6i.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c6id 인스턴스

c6id.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6id.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6id.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6id.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6id.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6id.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6id.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6id.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c6in 인스턴스

c6in.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c6in.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c6in.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c6in.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c6in.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c6in.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c6in.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c6in.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c7a 인스턴스

c7a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7a.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

c7a.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

c7g 인스턴스

c7g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7gd 인스턴스

c7gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7gn 인스턴스

c7gn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7gn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7gn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7gn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7gn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7gn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7i 인스턴스

c7i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c7i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c7i.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

c7i-flex 인스턴스

c7i-flex.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c7i-flex.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c7i-flex.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c7i-flex.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c7i-flex.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c7i-flex.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8g 인스턴스

c8g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8g.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8g.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

c8gd 인스턴스

c8gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2124
YARN_PROXYSERVER_HEAPSIZE	2124
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2124
HADOOP_NAMENODE_HEAPSIZE	972
HADOOP_DATANODE_HEAPSIZE	588

c8gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

c8gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

c8gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

c8gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

c8gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

c8gd.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

c8gd.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

d2 인스턴스

d2.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d2.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d2.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d2.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

d3 인스턴스

d3.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

d3en 인스턴스

d3en.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

d3en.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

d3en.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

d3en.6xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

d3en.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

d3en.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

f2 인스턴스

f2.6xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

f2.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

f2.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

g3 인스턴스

g3.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g3.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g3.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

g3s 인스턴스

g3s.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g4dn 인스턴스

g4dn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

g4dn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

g4dn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

g4dn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

g4dn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

g4dn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

g5 인스턴스

g5.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

g5.2xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g5.4xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g5.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g5.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

g5.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g5.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g5.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

g6 인스턴스

g6.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

g6.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

g6.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3317
YARN_PROXYSERVER_HEAPSIZE	3317
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3317
HADOOP_NAMENODE_HEAPSIZE	6348
HADOOP_DATANODE_HEAPSIZE	1679

g6.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

g6.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5795
YARN_PROXYSERVER_HEAPSIZE	5795
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5795
HADOOP_NAMENODE_HEAPSIZE	18739
HADOOP_DATANODE_HEAPSIZE	4096

g6.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

g6.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9543
YARN_PROXYSERVER_HEAPSIZE	9543
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9543
HADOOP_NAMENODE_HEAPSIZE	37478
HADOOP_DATANODE_HEAPSIZE	4096

g6.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	16998
YARN_PROXYSERVER_HEAPSIZE	16998
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	16998
HADOOP_NAMENODE_HEAPSIZE	74752
HADOOP_DATANODE_HEAPSIZE	4096

g6e 인스턴스

g6e.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

g6e.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

g6e.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

g6e.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

g6e.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

g6e.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

g6e.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

g6e.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

gr6 인스턴스

gr6.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4567
YARN_PROXYSERVER_HEAPSIZE	4567
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4567
HADOOP_NAMENODE_HEAPSIZE	12595
HADOOP_DATANODE_HEAPSIZE	2928

gr6.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7045
YARN_PROXYSERVER_HEAPSIZE	7045
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7045
HADOOP_NAMENODE_HEAPSIZE	24985
HADOOP_DATANODE_HEAPSIZE	4096

h1 인스턴스

h1.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

h1.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

h1.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

h1.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

i2 인스턴스

i2.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i2.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i2.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i2.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i3 인스턴스

i3.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i3.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i3.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i3.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i3.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i3en 인스턴스

i3en.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

i3en.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

i3en.3xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

i3en.6xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

i3en.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

i3en.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

i4g 인스턴스

i4g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i4i 인스턴스

i4i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i4i.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i4i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i4i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i4i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i4i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i4i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i4i.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

i7i 인스턴스

i7i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7i.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i7i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i7i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i7i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7i.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

i7ie 인스턴스

i7ie.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i7ie.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i7ie.3xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

i7ie.6xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.18xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	13332
YARN_PROXYSERVER_HEAPSIZE	13332
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	13332
HADOOP_NAMENODE_HEAPSIZE	56422
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i7ie.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

i8g 인스턴스

i8g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

i8g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

i8g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

i8g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

i8g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

i8g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

i8g.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

i8g.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

im4gn 인스턴스

im4gn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

im4gn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

im4gn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

im4gn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

im4gn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

is4gen 인스턴스

is4gen.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2557
YARN_PROXYSERVER_HEAPSIZE	2557
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2557
HADOOP_NAMENODE_HEAPSIZE	2547
HADOOP_DATANODE_HEAPSIZE	919

4gn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

is4gen.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

is4gen.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m1 인스턴스

m1.small

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	256
YARN_PROXYSERVER_HEAPSIZE	96
YARN_NODEMANAGER_HEAPSIZE	192
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	128
HADOOP_NAMENODE_HEAPSIZE	192
HADOOP_DATANODE_HEAPSIZE	96

m1.medium

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	384
YARN_PROXYSERVER_HEAPSIZE	192
YARN_NODEMANAGER_HEAPSIZE	256
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	256
HADOOP_NAMENODE_HEAPSIZE	384
HADOOP_DATANODE_HEAPSIZE	192

m1.large

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	768
YARN_PROXYSERVER_HEAPSIZE	384
YARN_NODEMANAGER_HEAPSIZE	512
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	512
HADOOP_NAMENODE_HEAPSIZE	768
HADOOP_DATANODE_HEAPSIZE	384

m1.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	1024
YARN_PROXYSERVER_HEAPSIZE	512
YARN_NODEMANAGER_HEAPSIZE	768
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	2304
HADOOP_DATANODE_HEAPSIZE	384

m2 인스턴스

m2.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1024
HADOOP_NAMENODE_HEAPSIZE	3072
HADOOP_DATANODE_HEAPSIZE	384

m2.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	1536
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1024
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	6144
HADOOP_DATANODE_HEAPSIZE	384

m2.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2048
YARN_PROXYSERVER_HEAPSIZE	1024
YARN_NODEMANAGER_HEAPSIZE	1536
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	1536
HADOOP_NAMENODE_HEAPSIZE	12288
HADOOP_DATANODE_HEAPSIZE	384

m3 인스턴스

m3.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m3.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2703
YARN_PROXYSERVER_HEAPSIZE	2703
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2703
HADOOP_NAMENODE_HEAPSIZE	3276
HADOOP_DATANODE_HEAPSIZE	1064

m4 인스턴스

m4.large

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2252
YARN_PROXYSERVER_HEAPSIZE	2252
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2252
HADOOP_NAMENODE_HEAPSIZE	1024
HADOOP_DATANODE_HEAPSIZE	614

m4.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m4.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m4.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m4.10xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5365
YARN_PROXYSERVER_HEAPSIZE	5365
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5365
HADOOP_NAMENODE_HEAPSIZE	16588
HADOOP_DATANODE_HEAPSIZE	3727

m4.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5 인스턴스

m5.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

m5a 인스턴스

m5a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

m5ad 인스턴스

m5ad.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5ad.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5ad.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5ad.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5ad.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5ad.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m5d 인스턴스

m5d.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2416
YARN_PROXYSERVER_HEAPSIZE	2416
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2416
HADOOP_NAMENODE_HEAPSIZE	1843
HADOOP_DATANODE_HEAPSIZE	778

m5d.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

m5d.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

m5d.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

m5d.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

m5d.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

m5d.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

m5dn 인스턴스

m5dn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5dn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5dn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5dn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5dn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5dn.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m5n 인스턴스

m5n.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m5n.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5n.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m5n.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m5n.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m5n.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m5n.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m5zn 인스턴스

m5zn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2396
YARN_PROXYSERVER_HEAPSIZE	2396
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2396
HADOOP_NAMENODE_HEAPSIZE	1740
HADOOP_DATANODE_HEAPSIZE	757

m5zn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m5zn.3xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3025
YARN_PROXYSERVER_HEAPSIZE	3025
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3025
HADOOP_NAMENODE_HEAPSIZE	4889
HADOOP_DATANODE_HEAPSIZE	1387

m5zn.6xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3962
YARN_PROXYSERVER_HEAPSIZE	3962
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3962
HADOOP_NAMENODE_HEAPSIZE	9574
HADOOP_DATANODE_HEAPSIZE	2324

m5zn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6a 인스턴스

m6a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6a.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m6a.48xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

m6g 인스턴스

m6g.xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6gd 인스턴스

m6gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6i 인스턴스

m6i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6i.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6i.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m6id 인스턴스

m6id.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6id.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6id.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6id.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6id.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6id.16xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6id.24xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6id.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m6idn 인스턴스

m6idn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6idn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6idn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6idn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6idn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6idn.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m6in 인스턴스

m6in.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m6in.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m6in.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m6in.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m6in.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m6in.16xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m6in.24xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m6in.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m7a 인스턴스

m7a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7a.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

m7a.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

m7g 인스턴스

m7g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7gd 인스턴스

m7gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7i 인스턴스

m7i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m7i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m7i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m7i.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

m7i-flex 인스턴스

m7i-flex.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m7i-flex.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m7i-flex.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m7i-flex.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m7i-flex.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m7i-flex.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8g 인스턴스

m8g.xLarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5877
YARN_PROXYSERVER_HEAPSIZE	5877
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5877
HADOOP_NAMENODE_HEAPSIZE	19148
HADOOP_DATANODE_HEAPSIZE	4096

m8g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8g.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m8g.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

m8gd 인스턴스

m8gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2401
YARN_PROXYSERVER_HEAPSIZE	2401
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2401
HADOOP_NAMENODE_HEAPSIZE	1766
HADOOP_DATANODE_HEAPSIZE	762

m8gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

m8gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

m8gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

m8gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	5836
YARN_PROXYSERVER_HEAPSIZE	5836
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	5836
HADOOP_NAMENODE_HEAPSIZE	18944
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

m8gd.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

p2 인스턴스

p2.xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p2.8xlarge

파라미터	값
YARN_RESOURCEMANAG ER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORY SERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

p2.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

p3 인스턴스

p3.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

p3.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

p3.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

p4d 인스턴스

p4d.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	24576
YARN_PROXYSERVER_HEAPSIZE	24576
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	24576
HADOOP_NAMENODE_HEAPSIZE	112640
HADOOP_DATANODE_HEAPSIZE	4096

p5 인스턴스

p5.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

r3 인스턴스

r3.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r3.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r3.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r3.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r4 인스턴스

r4.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r4.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r4.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r4.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r4.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5 인스턴스

r5.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

r5a 인스턴스

r5a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

r5ad 인스턴스

r5ad.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5ad.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5ad.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5ad.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12247
YARN_PROXYSERVER_HEAPSIZE	12247
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12247
HADOOP_NAMENODE_HEAPSIZE	50995
HADOOP_DATANODE_HEAPSIZE	4096

r5ad.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r5b 인스턴스

r5b.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5b.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5b.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5b.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5b.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5b.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5b.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r5d 인스턴스

r5d.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

r5d.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

r5d.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4710
YARN_PROXYSERVER_HEAPSIZE	4710
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4710
HADOOP_NAMENODE_HEAPSIZE	13312
HADOOP_DATANODE_HEAPSIZE	3072

r5d.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7331
YARN_PROXYSERVER_HEAPSIZE	7331
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7331
HADOOP_NAMENODE_HEAPSIZE	26419
HADOOP_DATANODE_HEAPSIZE	4096

r5d.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

r5d.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12574
YARN_PROXYSERVER_HEAPSIZE	12574
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12574
HADOOP_NAMENODE_HEAPSIZE	52633
HADOOP_DATANODE_HEAPSIZE	4096

r5d.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17817
YARN_PROXYSERVER_HEAPSIZE	17817
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17817
HADOOP_NAMENODE_HEAPSIZE	78848
HADOOP_DATANODE_HEAPSIZE	4096

r5dn 인스턴스

r5dn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5dn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5dn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5dn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5dn.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r5n 인스턴스

r5n.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r5n.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r5n.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r5n.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r5n.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r5n.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r5n.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6a 인스턴스

r6a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6a.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r6a.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

r6g 인스턴스

r6g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6gd 인스턴스

r6gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6i 인스턴스

r6i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6i.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

r6id 인스턴스

r6id.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6id.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6id.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6id.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6id.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6id.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6id.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6id.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r6idn 인스턴스

r6idn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6idn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6idn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6idn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6idn.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r6in 인스턴스

r6in.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r6in.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r6in.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r6in.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r6in.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r6in.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r6in.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r6in.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r7a 인스턴스

r7a.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7a.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7a.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7a.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7a.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7a.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7a.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7a.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

r7a.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

r7g 인스턴스

r7g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7gd 인스턴스

r7gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7i 인스턴스

r7i.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7i.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7i.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7i.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7i.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7i.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7i.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r7i.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

r7iz 인스턴스

r7iz.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r7iz.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r7iz.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r7iz.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r7iz.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

r8g 인스턴스

r8g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8g.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r8g.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

r8gd 인스턴스

r8gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2713
YARN_PROXYSERVER_HEAPSIZE	2713
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2713
HADOOP_NAMENODE_HEAPSIZE	3328
HADOOP_DATANODE_HEAPSIZE	1075

r8gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

r8gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

r8gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9584
YARN_PROXYSERVER_HEAPSIZE	9584
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9584
HADOOP_NAMENODE_HEAPSIZE	37683
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

r8gd.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x1 인스턴스

x1.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

x1e 인스턴스

x1e.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4520
YARN_PROXYSERVER_HEAPSIZE	4520
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4520
HADOOP_NAMENODE_HEAPSIZE	12364
HADOOP_DATANODE_HEAPSIZE	2882

x1e.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6952
YARN_PROXYSERVER_HEAPSIZE	6952
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6952
HADOOP_NAMENODE_HEAPSIZE	24524
HADOOP_DATANODE_HEAPSIZE	4096

x1e.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	11816
YARN_PROXYSERVER_HEAPSIZE	11816
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	11816
HADOOP_NAMENODE_HEAPSIZE	48844
HADOOP_DATANODE_HEAPSIZE	4096

x1e.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	21544
YARN_PROXYSERVER_HEAPSIZE	21544
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	21544
HADOOP_NAMENODE_HEAPSIZE	97484
HADOOP_DATANODE_HEAPSIZE	4096

x1e.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	41000
YARN_PROXYSERVER_HEAPSIZE	41000
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	41000
HADOOP_NAMENODE_HEAPSIZE	194764
HADOOP_DATANODE_HEAPSIZE	4096

x1e.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	79912
YARN_PROXYSERVER_HEAPSIZE	79912
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	79912
HADOOP_NAMENODE_HEAPSIZE	389324
HADOOP_DATANODE_HEAPSIZE	4096

x2gd 인스턴스

x2gd.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x2gd.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2gd.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x2gd.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2idn 인스턴스

x2idn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x2idn.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn 인스턴스

x2iedn.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x2iedn.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	42065
YARN_PROXYSERVER_HEAPSIZE	42065
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	42065
HADOOP_NAMENODE_HEAPSIZE	200089
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

x2iedn.32xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	82042
YARN_PROXYSERVER_HEAPSIZE	82042
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	82042
HADOOP_NAMENODE_HEAPSIZE	399974
HADOOP_DATANODE_HEAPSIZE	4096

x8g 인스턴스

x8g.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3338
YARN_PROXYSERVER_HEAPSIZE	3338
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3338
HADOOP_NAMENODE_HEAPSIZE	6451
HADOOP_DATANODE_HEAPSIZE	1699

x8g.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4587
YARN_PROXYSERVER_HEAPSIZE	4587
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4587
HADOOP_NAMENODE_HEAPSIZE	12697
HADOOP_DATANODE_HEAPSIZE	2949

x8g.4xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	7086
YARN_PROXYSERVER_HEAPSIZE	7086
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	7086
HADOOP_NAMENODE_HEAPSIZE	25190
HADOOP_DATANODE_HEAPSIZE	4096

x8g.8xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	12083
YARN_PROXYSERVER_HEAPSIZE	12083
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	12083
HADOOP_NAMENODE_HEAPSIZE	50176
HADOOP_DATANODE_HEAPSIZE	4096

x8g.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	17080
YARN_PROXYSERVER_HEAPSIZE	17080
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	17080
HADOOP_NAMENODE_HEAPSIZE	75161
HADOOP_DATANODE_HEAPSIZE	4096

x8g.16xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	22077
YARN_PROXYSERVER_HEAPSIZE	22077
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	22077
HADOOP_NAMENODE_HEAPSIZE	100147
HADOOP_DATANODE_HEAPSIZE	4096

x8g.24xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	32071
YARN_PROXYSERVER_HEAPSIZE	32071
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	32071
HADOOP_NAMENODE_HEAPSIZE	150118
HADOOP_DATANODE_HEAPSIZE	4096

x8g.48xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	62054
YARN_PROXYSERVER_HEAPSIZE	62054
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	62054
HADOOP_NAMENODE_HEAPSIZE	300032
HADOOP_DATANODE_HEAPSIZE	4096

z1d 인스턴스

z1d.xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	2744
YARN_PROXYSERVER_HEAPSIZE	2744
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	2744
HADOOP_NAMENODE_HEAPSIZE	3481
HADOOP_DATANODE_HEAPSIZE	1105

z1d.2xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	3399
YARN_PROXYSERVER_HEAPSIZE	3399
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	3399
HADOOP_NAMENODE_HEAPSIZE	6758
HADOOP_DATANODE_HEAPSIZE	1761

z1d.3xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	4055
YARN_PROXYSERVER_HEAPSIZE	4055
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	4055
HADOOP_NAMENODE_HEAPSIZE	10035
HADOOP_DATANODE_HEAPSIZE	2416

z1d.6xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	6021
YARN_PROXYSERVER_HEAPSIZE	6021
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	6021
HADOOP_NAMENODE_HEAPSIZE	19865
HADOOP_DATANODE_HEAPSIZE	4096

z1d.12xlarge

파라미터	값
YARN_RESOURCEMANAGER_HEAPSIZE	9953
YARN_PROXYSERVER_HEAPSIZE	9953
YARN_NODEMANAGER_HEAPSIZE	2048
HADOOP_JOB_HISTORYSERVER_HEAPSIZE	9953
HADOOP_NAMENODE_HEAPSIZE	39526
HADOOP_DATANODE_HEAPSIZE	4096

HDFS 구성

다음 표에서는 기본 하둡 분산 파일 시스템(HDFS) 파라미터 및 관련 설정을 설명합니다. hdfs-site 구성 분류를 사용하여 이러한 값을 변경할 수 있습니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

Warning

1. 노드 수가 4개 미만인 클러스터에서 `dfs.replication`을 1로 설정하면 단일 노드가 중단된 경우 HDFS 데이터가 손실될 수 있습니다. 클러스터에 HDFS 스토리지가 있는 경우 데이터 손실을 방지하려면 프로덕션 워크로드에 사용할 코어 노드를 4개 이상 포함하는 클러스터를 구성하는 것이 좋습니다.
2. Amazon EMR은 클러스터에서 코어 노드를 `dfs.replication` 미만으로 조정하는 허용하지 않습니다. 예를 들어, `dfs.replication = 2`인 경우 최소 코어 노드 수가 2개입니다.
3. Managed Scaling, Auto Scaling을 사용하거나 클러스터 크기를 수동으로 조정하는 경우 `dfs.replication`을 2 이상으로 설정하는 것이 좋습니다.

파라미터	정의	기본값
`dfs.block.size`	HDFS 블록의 크기입니다. HDFS에 저장된 데이터에 대해 작동하는 경우 분할 크기는 일반적으로 HDFS 블록의 크기입니다. 숫자가 클수록 작업 세분화가 적지만 NameNode 클러스터에 부담을 덜 줍니다.	134217728 (128MB)
`dfs.replication`	내구성을 위해 저장할 각 블록의 사본 수입니다. Amazon EMR은 클러스터에 프로비저닝된 코어 노드 수를 기반으로 이 값을 설정합니다. 필요에 맞게 값을 조정합니다. 기본값을 덮어쓰려면 `hdfs-site` 분류를 사용하십시오.	코어 노드가 4개 미만으로 프로비저닝된 클러스터의 경우 1 코어 노드가 10개 미만으로 프로비저닝된 클러스터의 경우 2

파라미터	정의	기본값
		기타 모든 클러스터의 경우 3

Amazon EMR에서 HDFS의 투명한 암호화

투명한 암호화는 사용자가 정의하는 HDFS 경로인 HDFS 암호화 영역의 사용을 통해 구현됩니다. 각 암호화 영역에는 고유의 키가 있으며, 이 키는 `hdfs-site` 구성 분류를 사용하여 지정된 키 서버에 저장됩니다.

Amazon EMR 릴리스 버전 4.8.0부터 Amazon EMR 보안 구성을 사용하여 클러스터의 데이터 암호화 설정을 더욱 쉽게 구성할 수 있습니다. 보안 구성은 Amazon S3의 EMRFS 데이터 및 Amazon Elastic Block Store(Amazon EBS) 스토리지 볼륨의 전송 중 데이터와 저장 데이터에 대한 보안을 활성화하는 설정을 제공합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 전송 중 데이터 및 저장 데이터 암호화를 참조하세요.

Amazon EMR은 기본적으로 Hadoop KMS를 사용하지만, KeyProvider API 작업을 구현하는 다른 KMS를 사용할 수 있습니다. HDFS 암호화 영역의 각 파일에는 고유의 데이터 암호화 키가 있으며 이 키는 암호와 영역 키를 사용하여 암호화됩니다. 암호화 및 암호 해독 활동은 클라이언트에서만 수행되므로 데이터를 암호화 영역에 쓸 때 HDFS 데이터가 종단 간(미사용 데이터 및 전송 중 데이터)으로 암호화됩니다.

암호화 영역 간에 또는 암호화 영역에서 암호화되지 않은 경로로 파일을 이동할 수 없습니다.

NameNode와 HDFS 클라이언트는 KeyProvider API 작업을 통해 하둡 KMS(또는 사용자가 구성한 대체 KMS)와 상호 작용합니다. KMS는 지원 키 스토어에 암호화 키를 저장하는 작업을 담당합니다. 또한 Amazon EMR에는 JCE 무제한 강도 정책이 포함되므로, 원하는 길이로 키를 생성할 수 있습니다.

자세한 내용은 Hadoop 설명서에서 Transparent encryption in HDFS를 참조하세요.

> **Note**
>
> Amazon EMR에서 Hadoop KMS를 사용할 경우 HTTPS를 통한 KMS는 기본적으로 활성화되지 않습니다. HTTPS를 통한 KMS를 활성화하는 방법에 대한 자세한 내용은 하둡 KMS 설명서를 참조하십시오.

HDFS 투명한 암호화 구성

키를 생성하고 암호화 영역을 추가하여 Amazon EMR에서 투명한 암호화를 구성할 수 있습니다. 여러 가지 방법으로 이 작업을 수행할 수 있습니다.

- 클러스터를 생성할 때 Amazon EMR 구성 API 작업 사용
- command-runner.jar과 함께 하둡 JAR 단계 사용
- 하둡 클러스터의 마스터 노드에 로그인하여 `hadoop key` 및 `hdfs crypto` 명령줄 클라이언트 사용
- 하둡 KMS 및 HDFS에 REST API 사용

REST API에 대한 자세한 내용은 하둡 KMS 및 HDFS의 각 설명서를 참조하십시오.

CLI를 사용하여 클러스터 생성 시 암호화 영역과 해당 키를 생성하려면

구성 API 작업의 `hdfs-encryption-zones` 분류를 통해 클러스터를 생성할 때 키 이름과 암호화 영역을 지정할 수 있습니다. Amazon EMR은 클러스터의 Hadoop KMS에서 이 키를 생성하고 암호화 영역을 구성합니다.

- 다음 명령을 사용하여 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`myConfig.json`:

```
[
 {
  "Classification": "hdfs-encryption-zones",
  "Properties": {
```

```
    "/myHDFSPath1": "path1_key",
    "/myHDFSPath2": "path2_key"
   }
  }
]
```

마스터 노드에서 암호화 영역과 해당 키를 수동으로 생성하려면

1. 4.1.0 이상의 Amazon EMR 릴리스를 사용하여 클러스터를 시작합니다.
2. SSH를 사용하여 클러스터의 프라이머리 노드에 연결합니다.
3. 하둡 KMS 내에서 키를 생성합니다.

```
$ hadoop key create path2_key
path2_key has been successfully created with options Options{cipher='AES/CTR/
NoPadding', bitLength=256, description='null', attributes=null}.
KMSClientProvider[http://ip-x-x-x-x.ec2.internal:16000/kms/v1/] has been updated.
```

> **Important**
>
> 하둡 KMS에서는 키 이름이 소문자여야 합니다. 대문자가 있는 키를 사용하면 클러스터가 시작 중 실패합니다.

4. HDFS에서 암호화 영역 경로를 생성합니다.

```
$ hadoop fs -mkdir /myHDFSPath2
```

5. 생성한 키를 사용하여 HDFS 경로를 암호화 영역으로 만듭니다.

```
$ hdfs crypto -createZone -keyName path2_key -path /myHDFSPath2
Added encryption zone /myHDFSPath2
```

를 사용하여 암호화 영역 및 해당 키를 수동으로 생성하려면 AWS CLI

• 다음 명령을 사용하여 KMS 키와 암호화 영역을 수동으로 생성하는 단계를 추가합니다.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Create
 First Hadoop KMS Key",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[/
bin/bash,-c,"\"hadoop key create path1_key\""] \
```

```
Type=CUSTOM_JAR,Name="Create First Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create First Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path1_key -path /myHDFSPath1\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop KMS Key",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop key create
 path2_key\""] \
Type=CUSTOM_JAR,Name="Create Second Hadoop HDFS Path",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hadoop fs -mkdir /
myHDFSPath2\""] \
Type=CUSTOM_JAR,Name="Create Second Encryption Zone",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[/bin/bash,-c,"\"hdfs crypto -createZone
 -keyName path2_key -path /myHDFSPath2\""]
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

HDFS 투명한 암호화에 대한 고려 사항

모범 사례는 파일을 작성할 수 있는 각 애플리케이션에 대해 암호화 영역을 생성하는 것입니다. 또한 구성 API에서 hdfs-encryption-zones 분류를 사용하여 모든 HDFS를 암호화하고 루트 경로(/)를 암호화 영역으로 지정할 수 있습니다.

Hadoop 키 관리 서버

하둡 KMS는 하둡 클러스터에 대한 암호화 서비스를 구현할 수 있는 기능을 제공하는 키 관리 서버이며, Amazon EMR에서 HDFS의 투명한 암호화 항목의 키 공급자 역할을 수행할 수 있습니다. EMR 클러스터를 시작하는 동안 Hadoop 애플리케이션을 선택하면 Amazon EMR의 Hadoop KMS가 기본적으로 설치되고 활성화됩니다. 하둡 KMS는 임시 캐시의 경우를 제외하고 키 자체를 저장하지 않습니다. Hadoop KMS는 키 제공업체 및 지원 키 스토어에 대한 클라이언트 수탁자 사이에서 프록시 역할을 하지만, 키 스토어는 아닙니다. 하둡 KMS에 대해 생성되는 기본 키 스토어는 JCEKS(Java Cryptography Extension KeyStore)입니다. JCE 무제한 강도 정책도 포함되므로, 원하는 길이로 키를 생성할 수 있습니다. 또한 하둡 KMS는 HDFS와 같은 다른 클라이언트 애플리케이션과 독립적으로 키 및 키 작업에 대한 액세스를 제어하는 다양한 ACL을 지원합니다. Amazon EMR의 기본 키 길이는 256비트입니다.

하둡 KMS를 구성하려면 hadoop-kms-site 분류를 사용하여 설정을 변경합니다. ACL을 구성하려면 kms-acls 분류를 사용합니다.

자세한 내용은 [하둡 KMS 설명서]를 참조하십시오. 하둡 KMS는 하둡 HDFS 투명한 암호화에 사용됩니다. HDFS 투명한 암호화에 대해 자세히 알아보려면 Apache Hadoop 설명서의 [HDFS transparent encryption] 주제를 참조하세요.

> **Note**
>
> Amazon EMR에서 Hadoop KMS를 사용할 경우 HTTPS를 통한 KMS는 기본적으로 활성화되지 않습니다. HTTPS를 통한 KMS를 활성화하는 방법을 알아보려면 [하둡 KMS 설명서]를 참조하십시오.

> **Important**
>
> 하둡 KMS에서는 키 이름이 소문자여야 합니다. 대문자가 있는 키를 사용하면 클러스터가 시작 중 실패합니다.

Amazon EMR에서 Hadoop KMS 구성

Amazon EMR 릴리스 버전 4.6.0 이상을 사용할 경우 `kms-http-port`는 9700이며 `kms-admin-port`는 9701입니다.

Amazon EMR 릴리스에 대한 구성 API를 사용하여 클러스터 생성 시 Hadoop KMS를 구성할 수 있습니다. 다음은 하둡 KMS에 사용할 수 있는 구성 객체 분류입니다.

Hadoop KMS 구성 분류

분류	파일 이름
hadoop-kms-site	`kms-site.xml`
hadoop-kms-acls	`kms-acls.xml`
hadoop-kms-env	`kms-env.sh`
hadoop-kms-log4j	`kms-log4j.properties`

CLI를 사용하여 하둡 KMS ACL를 설정하려면

- 다음 명령을 사용하여 ACL이 포함된 하둡 KMS에 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-acls",
  "Properties": {
   "hadoop.kms.blacklist.CREATE": "hdfs,foo,myBannedUser",
   "hadoop.kms.acl.ROLLOVER": "myAllowedUser"
  }
 }
]
```

CLI를 사용하여 하둡 KMS를 비활성화하려면

- 다음 명령을 사용하여 하둡 KMS `hadoop.kms.cache.enable`이 `false`로 설정된 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-site",
  "Properties": {
   "hadoop.kms.cache.enable": "false"
  }
 }
]
```

CLI를 사용하여 **kms-env.sh** 스크립트에서 환경 변수를 설정하려면

- `kms-env.sh`에서 hadoop-kms-env 구성을 통해 설정을 변경합니다. 다음 명령을 사용하여 하둡 KMS가 포함된 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=App1 Name=App2 --configurations https://s3.amazonaws.com/amzn-
s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`myConfig.json`:

```
[
 {
  "Classification": "hadoop-kms-env",
```

```
    "Properties": {},
    "Configurations": [
     {
      "Classification": "export",
      "Properties": {
       "JAVA_LIBRARY_PATH": "/path/to/files",
       "KMS_SSL_KEYSTORE_FILE": "/non/Default/Path/.keystore",
       "KMS_SSL_KEYSTORE_PASS": "myPass"
      },
      "Configurations": []
     }
    ]
  }
]
```

하둡 KMS 구성에 대한 자세한 내용은 하둡 KMS 설명서를 참조하십시오.

여러 프라이머리 노드가 있는 EMR 클러스터에서 HDFS의 투명한 암호화

Apache Ranger KMS는 HDFS의 투명한 암호화를 위해 여러 프라이머리 노드가 있는 Amazon EMR 클러스터에서 사용됩니다.

Apache Ranger KMS는 여러 프라이머리 노드가 있는 Amazon EMR 클러스터의 루트 키와 암호화 영역(EZ) 키를 Amazon RDS에 저장합니다. 여러 프라이머리가 있는 Amazon EMR 클러스터에서 HDFS의 투명한 암호화를 활성화하려면 다음 구성을 제공해야 합니다.

- Ranger KMS 루트 키 및 EZ 키를 저장하기 위한 Amazon RDS 또는 자체 MySQL 서버 연결 URL
- MySQL의 사용자 이름 및 암호
- Ranger KMS 루트 키의 암호
- MySQL 서버에 대한 SSL 연결 시 사용할 인증 기관(CA) PEM 파일. Amazon RDS용 인증서 번들 다운로드 AWS 리전 에서 용 인증서 번들을 다운로드할 수 있습니다. https://docs.aws.amazon.com/AmazonRDS/latest/UserGuide/UsingWithRDS.SSL.html#UsingWithRDS.SSL.CertificatesDownload

다음 예제에서 보여주듯이 `ranger-kms-dbks-site` 분류 및 `ranger-kms-db-ca` 분류를 사용하여 이러한 구성을 제공할 수 있습니다.

```
[
 {
  "Classification": "ranger-kms-dbks-site",
```

```
  "Properties": {
   "ranger.ks.jpa.jdbc.url": "jdbc:log4jdbc:mysql://mysql-host-url.xx-
xxx-1.xxx.amazonaws.com:3306/rangerkms",
   "ranger.ks.jpa.jdbc.user": "mysql-user-name",
   "ranger.ks.jpa.jdbc.password": "mysql-password",
   "ranger.db.encrypt.key.password": "password-for-encrypting-a-master-key"
  }
 },
 {
  "Classification": "ranger-kms-db-ca",
  "Properties": {
   "ranger.kms.trust.ca.file.s3.url": "<S3-path-of-downloaded-pem-file>"
  }
 }
]
```

다음은 Apache Ranger KMS를 위한 구성 객체 분류입니다.

Hadoop KMS 구성 분류

분류	설명
ranger-kms-dbks-site	Ranger KMS에서 dbks-site.xml 파일의 값을 변경합니다.
ranger-kms-site	Ranger KMS에서 ranger-kms-site.xml 파일의 값을 변경합니다.
ranger-kms-env	Ranger KMS 환경의 값을 변경합니다.
ranger-kms-log4j	Ranger KM에서 kms-log4j.properties 파일의 값을 변경합니다.
ranger-kms-db-ca	Ranger KMS와의 MySQL SSL 연결에서 CA 파일의 값을 변경합니다.

고려 사항

- 보안을 향상시키려면 Amazon RDS 인스턴스를 암호화하는 것이 좋습니다. 자세한 내용은 Amazon RDS 리소스 암호화 개요를 참조하세요.

- 보안 수준을 높이려면 여러 프라이머리 노드가 있는 각 Amazon EMR 클러스터에 대해 별도의 MySQL 데이터베이스를 사용하는 것이 좋습니다.
- 여러 프라이머리 노드가 있는 Amazon EMR 클러스터에서 HDFS의 투명한 암호를 구성하려면 클러스터를 생성할 때 `hdfs-encryption-zones` 분류를 지정해야 합니다. 그렇지 않으면 Ranger KMS는 구성되거나 시작되지 않습니다. 실행 중인 클러스터에서의 모든 Hadoop KMS 구성 분류 또는 `hdfs-encryption-zones` 분류의 재구성은 여러 프라이머리 노드가 있는 Amazon EMR 클러스터에서 지원되지 않습니다.
- Amazon RDS용 인증서 번들 다운로드에서 다운로드한 PEM 인증서 번들은 여러 인증서를 하나의 파일로 그룹화합니다. Amazon EMR 7.3.0 이상에서는 `ranger.kms.trust.ca.file.s3.url` 구성을 사용하여 PEM 파일에서 여러 인증서 가져오기를 지원합니다.

Hadoop 애플리케이션 생성 또는 실행

주제

- Amazon EMR을 사용하여 바이너리 빌드
- 스트리밍을 사용하여 데이터 처리
- 사용자 지정 JAR로 데이터 처리

Amazon EMR을 사용하여 바이너리 빌드

Amazon EMR을 빌드 환경으로 사용하여 클러스터에서 사용할 수 있도록 프로그램을 컴파일할 수 있습니다. Amazon EMR에서 사용하는 프로그램은 Amazon EMR에 사용되는 것과 동일한 Linux 버전을 실행하는 시스템에서 컴파일해야 합니다. 32비트 버전의 경우 32비트 시스템에서 컴파일하거나 32비트 교차 컴파일 옵션이 활성화된 상태로 컴파일해야 합니다. 64비트 버전의 경우 64비트 시스템에서 컴파일하거나 64비트 교차 컴파일 옵션이 활성화된 상태로 컴파일해야 합니다. EC2 인스턴스 버전에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 EC2 인스턴스 계획 및 구성을 참조하세요. 지원되는 프로그래밍 언어로는 C++, Python 및 C#이 있습니다.

다음 테이블에는 Amazon EMR을 사용한 애플리케이션 빌드 및 테스트 관련 단계가 요약되어 있습니다.

모듈 빌드 프로세스

1	클러스터의 마스터 노드에 연결합니다.
2	소스 파일을 마스터 노드로 복사합니다.

3	필요한 최적화를 사용하여 바이너리를 빌드합니다.
4	프라이머리 노드에서 Amazon S3으로 바이너리를 복사합니다.

이러한 각각의 단계에 대한 자세한 내용은 다음 단원에서 설명합니다.

클러스터의 마스터 노드에 연결하려면

- Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결의 지침을 따릅니다.

소스 파일을 마스터 노드로 복사하려면

1. Amazon S3 버킷에 소스 파일을 배치합니다. 버킷을 생성하는 방법과 Amazon S3로 데이터를 이동하는 방법을 알아보려면 Amazon Simple Storage Service 사용 설명서를 참조하세요.
2. 다음과 비슷한 명령을 입력하여 하둡 클러스터에 소스 파일용 폴더를 생성합니다.

   ```
   mkdir SourceFiles
   ```
3. 다음과 비슷한 명령을 입력하여 Amazon S3에서 프라이머리 노드로 소스 파일을 복사합니다.

   ```
   hadoop fs -get s3://amzn-s3-demo-bucket/SourceFiles SourceFiles
   ```

필요한 최적화를 사용하여 바이너리 빌드

바이너리 빌드 방식은 여러 요인에 따라 달라집니다. 특정 빌드 도구에 대한 지침에 따라 환경을 설정하고 구성합니다. 빌드 환경을 설치하는 방법을 알아보기 위해 하둡 시스템 사양 명령을 사용하여 클러스터 정보를 가져올 수 있습니다.

시스템 사양을 식별하려면

- 다음 명령을 사용하여 바이너리를 빌드하는 데 사용하려는 아키텍처를 확인합니다.

  a. Debian 버전을 보려면 다음 명령을 입력합니다.

     ```
     master$ cat /etc/issue
     ```

     출력 결과는 다음과 비슷합니다.

```
Debian GNU/Linux 5.0
```

b. 퍼블릭 DNS 이름 및 프로세서 크기를 보려면 다음 명령을 입력합니다.

```
master$ uname -a
```

출력 결과는 다음과 비슷합니다.

```
Linux domU-12-31-39-17-29-39.compute-1.internal 2.6.21.7-2.fc8xen #1 SMP Fri
 Feb 15 12:34:28 EST 2008 x86_64 GNU/Linux
```

c. 프로세서 속도를 보려면 다음 명령을 입력합니다.

```
master$ cat /proc/cpuinfo
```

출력 결과는 다음과 비슷합니다.

```
processor : 0
vendor_id : GenuineIntel
model name : Intel(R) Xeon(R) CPU E5430 @ 2.66GHz
flags : fpu tsc msr pae mce cx8 apic mca cmov pat pse36 clflush dts acpi mmx
 fxsr sse sse2 ss ht tm syscall nx lm constant_tsc pni monitor ds_cpl vmx est
 tm2 ssse3 cx16 xtpr cda lahf_lm
...
```

바이너리가 빌드되면 파일을 Amazon S3로 복사할 수 있습니다.

프라이머리 노드에서 Amazon S3로 바이너리를 복사하는 방법

- 다음 명령을 입력하여 바이너리를 Amazon S3 버킷으로 복사합니다.

```
hadoop fs -put BinaryFiles s3://amzn-s3-demo-bucket/BinaryDestination
```

스트리밍을 사용하여 데이터 처리

Hadoop Streaming은 Hadoop과 함께 제공되는 유틸리티로, 사용자는 이 유틸리티를 이용해 Java 이외의 언어로 MapReduce 실행 파일을 개발할 수 있습니다. 스트리밍은 JAR 파일 형식으로 구현되므로 Amazon EMR API 또는 명령줄에서 표준 JAR 파일과 같이 스트리밍을 실행할 수 있습니다.

이 섹션에서는 Amazon EMR에서 스트리밍을 사용하는 방법을 설명합니다.

> **Note**
>
> Apache Hadoop 스트리밍은 독립적인 도구입니다. 따라서 모든 기능 및 파라미터는 여기에 설명되어 있지 않습니다. Hadoop 스트리밍에 대한 자세한 내용은 http://hadoop.apache.org/docs/stable/hadoop-streaming/HadoopStreaming.html을 참조하세요.

Hadoop 스트리밍 유틸리티 사용

이 섹션에서는 Hadoop의 스트리밍 유틸리티를 사용하는 방법을 설명합니다.

Hadoop 프로세스

1	원하는 프로그래밍 언어로 mapper 및 reducer 실행 파일을 작성하십시오. Hadoop 설명서의 지침에 따라 스트리밍 실행 파일을 작성하십시오. 프로그램은 표준 출력을 통해 표준 입력 및 출력 데이터에서 입력을 읽어야 합니다. 기본적으로 각 입력/출력 행은 레코드를 나타내며 각 행의 첫 번째 탭은 키와 값 사이의 구분 기호로 사용됩니다.
2	실행 파일을 로컬에서 테스트하여 Amazon S3에 업로드합니다.
3	Amazon EMR 명령줄 인터페이스 또는 Amazon EMR 콘솔을 사용하여 애플리케이션을 실행합니다.

각 mapper 스크립트는 클러스터에서 별도의 프로세스로 실행됩니다. 각 reducer 실행 파일은 mapper 실행 파일의 출력을 작업 흐름에 의해 데이터 출력으로 변환합니다.

대부분의 스트리밍 애플리케이션에는 `input`, `output`, `mapper` 및 `reducer` 파라미터가 필요합니다. 다음 표에서는 이 파라미터와 다른 선택적 파라미터를 설명합니다.

파라미터	설명	필수
-입력	Amazon S3에서 입력 데이터의 위치. 유형: 문자열 기본값: None 제약 조건: URI 프로토콜을 지정하지 않으면 클러스터의 기본 파일 시스템을 사용합니다.	예
-출력	Amazon EMR이 처리된 데이터를 업로드하는 Amazon S3의 위치. 유형: 문자열 기본값: None 제약 조건: URI 기본값: 위치를 지정하지 않으면 Amazon EMR이 `input`에서 지정하는 위치로 데이터를 업로드합니다.	예
-mapper	mapper 실행 파일의 이름. 유형: 문자열 기본값: None	예
-reducer	reducer 실행 파일의 이름. 유형: 문자열 기본값: None	예
-cacheFile	주로 성능 향상을 위해 Hadoop이 로컬 작업 디렉터리에 복사할 파일이 들어있는 Amazon S3의 위치.	아니요

파라미터	설명	필수
	유형: 문자열 기본값: None 제약 조건: [URI]#[작업 디렉터리에 생성할 symlink 이름]	
-cacheArchive	작업 디렉터리에 추출할 JAR 파일 유형: 문자열 기본값: None 제약 조건: [URI]#[작업 디렉터리에 생성할 symlink 디렉터리 이름]	아니요
-combiner	결과를 결합 유형: 문자열 기본값: None 제약 조건: Java 클래스 이름	아니요

다음 코드 샘플은 Python으로 작성된 mapper 실행 파일입니다. 이 스크립트는 WordCount 샘플 애플리케이션의 일부입니다.

```
#!/usr/bin/python
import sys

def main(argv):
  line = sys.stdin.readline()
  try:
    while line:
      line = line.rstrip()
      words = line.split()
      for word in words:
        print "LongValueSum:" + word + "\t" + "1"
      line = sys.stdin.readline()
```

```
    except "end of file":
      return None
if __name__ == "__main__":
  main(sys.argv)
```

스트리밍 단계 제출

이 섹션에서는 클러스터에 스트리밍 단계를 제출하는 기본 사항을 설명합니다. 스트리밍 애플리케이션은 표준 입력에서 입력을 읽은 다음 각 입력에 대해 스크립트 또는 실행 파일(매퍼라고 함)을 실행합니다. 각 입력의 결과는 일반적으로 하둡 분산 파일 시스템(HDFS) 파티션에 로컬로 저장됩니다. 모든 입력이 mapper에 의해 처리된 후 두 번째 스크립트 또는 실행 파일(reducer라고 함)이 mapper 결과를 처리합니다. reducer의 결과는 표준 출력으로 전송됩니다. 한 단계의 출력이 다른 단계의 입력이 되는 일련의 스트리밍 단계를 함께 연결할 수 있습니다.

mapper와 reducer는 각각 파일로 참조되거나 사용자가 Java 클래스를 제공할 수 있습니다. Ruby, Perl, Python, PHP 또는 Bash를 비롯한 지원되는 언어로 mapper 및 reducer를 구현할 수 있습니다.

콘솔을 사용하여 스트리밍 단계 제출

이 예제는 Amazon EMR 콘솔을 사용하여 스트리밍 단계를 실행 중인 클러스터에 제출하는 방법을 설명합니다.

스트리밍 단계를 제출하는 방법

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 목록에서 클러스터의 이름을 선택합니다.
3. 단계 섹션으로 스크롤하여 섹션을 확장한 후 단계 추가를 선택합니다.
4. 단계 추가 대화 상자에서 다음을 수행합니다.

   - Step type(단계 유형)에서 Streaming program(스트리밍 프로그램)을 선택합니다.
   - Name(이름)에서 기본 이름(Streaming program)을 수락하거나 새 이름을 입력합니다.
   - Mapper(매퍼)에서 하둡의 매퍼 클래스 위치 또는 매퍼 실행 파일(예: Python 프로그램)이 있는 S3 버킷을 입력하거나 찾습니다. 경로 값은 *BucketName*/*path*/*MapperExecutable* 형식이어야 합니다.
   - Reducer에서 하둡의 reducer 클래스 위치 또는 reducer 실행 파일(예: Python 프로그램)이 있는 S3 버킷을 입력하거나 찾습니다. 경로 값은 *BucketName*/*path*/*MapperExecutable* 형식이어야 합니다. Amazon EMR은 특수 aggregate 키워드를 지원합니다. 자세한 내용은 Hadoop에서 제공하는 Aggregate 라이브러리를 참조하십시오.

- Input S3 location(입력 S3 위치)에서 입력 데이터의 위치를 입력하거나 찾아봅니다.
- 출력 S3 위치에서 Amazon S3 출력 버킷의 이름을 입력하거나 찾아봅니다.
- 인수에서 필드를 비워 둡니다.
- 실패 시 작업에서 기본 옵션(계속)을 그대로 사용합니다.

5. 추가를 선택합니다. 단계가 콘솔에 [Pending] 상태로 나타납니다.
6. 단계의 상태는 단계가 실행됨에 따라 대기 중에서 실행 중을 거쳐 완료됨으로 바뀝니다. 상태를 업데이트하려면 Actions(작업) 열 위의 Refresh(새로 고침) 아이콘을 선택합니다.

AWS CLI

이 예제에서는를 사용하여 클러스터 AWS CLI 를 생성하고 스트리밍 단계를 제출하는 방법을 보여줍니다.

를 사용하여 클러스터를 생성하고 스트리밍 단계를 제출하려면 AWS CLI

- 를 사용하여 클러스터를 생성하고 스트리밍 단계를 제출하려면 다음 명령을 AWS CLI입력하고 *myKey*를 EC2 키 페어의 이름으로 바꿉니다. `--files`에 대한 인수는 스크립트 위치에 대한 Amazon S3 경로여야 하고, `-mapper` 및 `-reducer`에 대한 인수는 해당 스크립트 파일의 이름이어야 합니다.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 --
applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=STREAMING,Name="Streaming Program",ActionOnFailure=CONTINUE,Args=[--
files,pathtoscripts,-mapper,mapperscript,-reducer,reducerscript,aggregate,-
input,pathtoinputdata,-output,pathtooutputbucket]
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`--instance-groups` 파라미터를 사용하지 않고 인스턴스 수를 지정하면 단일 마스터 노드가 시작되고 나머지 인스턴스는 코어 노드로 시작됩니다. 모든 노드에는 이 명령에 지정된 인스턴스 유형이 사용됩니다.

> **Note**
>
> Amazon EMR 서비스 역할과 EC2 인스턴스 프로파일을 아직 생성하지 않았다면 `emr create-default-roles` 하위 명령을 입력하기 전에 aws `create-cluster`를 입력하여 생성합니다.

에서 Amazon EMR 명령을 사용하는 방법에 대한 자세한 내용은 섹션을 AWS CLI참조하세요https://docs.aws.amazon.com/cli/latest/reference/emr.

사용자 지정 JAR로 데이터 처리

사용자 지정 JAR은 Amazon S3에 업로드할 수 있는 컴파일된 Java 프로그램을 실행합니다. 시작하려는 Hadoop의 버전에 맞게 프로그램을 컴파일하고 Amazon EMR `CUSTOM_JAR` 단계를 제출해야 합니다. JAR 파일 컴파일에 대한 자세한 내용은 [Amazon EMR을 사용하여 바이너리 빌드](#) 섹션을 참조하세요.

Hadoop MapReduce 애플리케이션 빌드 방법에 대한 자세한 내용은 Apache Hadoop 설명서에서 [MapReduce Tutorial](#)을 참조하세요.

주제

- [사용자 지정 JAR 단계 제출](#)

사용자 지정 JAR 단계 제출

사용자 지정 JAR은 Amazon S3에 업로드할 수 있는 컴파일된 Java 프로그램을 실행합니다. 시작하려는 Hadoop의 버전에 맞게 프로그램을 컴파일하고 Amazon EMR `CUSTOM_JAR` 단계를 제출해야 합니다. JAR 파일 컴파일에 대한 자세한 내용은 [Amazon EMR을 사용하여 바이너리 빌드](#) 섹션을 참조하세요.

Hadoop MapReduce 애플리케이션 빌드 방법에 대한 자세한 내용은 Apache Hadoop 설명서에서 [MapReduce Tutorial](#)을 참조하세요.

이 섹션에서는 Amazon EMR에서 사용자 지정 JAR 단계 제출의 기본 사항을 다룹니다. 사용자 지정 JAR 단계를 제출하면 Java 프로그래밍 언어를 사용하여 데이터를 처리하기 위한 스크립트를 작성할 수 있습니다.

콘솔을 사용하여 사용자 지정 JAR 단계 제출

이 예제는 Amazon EMR 콘솔을 사용하여 사용자 지정 JAR 단계를 실행 중인 클러스터에 제출하는 방법을 설명합니다.

콘솔을 사용하여 사용자 지정 JAR 단계를 제출하는 방법

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 목록에서 클러스터의 이름을 선택합니다.
3. 단계 섹션으로 스크롤하여 섹션을 확장한 후 단계 추가를 선택합니다.
4. 단계 추가 대화 상자에서 다음을 수행합니다.
   - Step type(단계 유형)에서 Custom JAR(사용자 지정 JAR)을 선택합니다.
   - Name(이름)에서 기본 이름(Custom JAR)을 수락하거나 새 이름을 입력합니다.
   - JAR S3 location(JAR S3 위치)에서 JAR 파일의 위치를 입력하거나 찾아봅니다. JAR 위치는 S3에 대한 경로 또는 클래스 경로의 정규화된 Java 클래스일 수 있습니다.
   - Arguments(인수)에 필수 인수를 공백으로 구분된 문자열로 입력하거나 필드를 비워 둡니다.
   - 실패 시 작업에서 기본 옵션(계속)을 그대로 사용합니다.
5. 추가를 선택합니다. 단계가 콘솔에 [Pending] 상태로 나타납니다.
6. 단계의 상태는 단계가 실행됨에 따라 대기 중에서 실행 중을 거쳐 완료됨으로 바뀝니다. 상태를 업데이트하려면 Actions(작업) 열 위의 Refresh(새로 고침) 아이콘을 선택합니다.

를 사용하여 클러스터 시작 및 사용자 지정 JAR 단계 제출 AWS CLI

클러스터를 시작하고를 사용하여 사용자 지정 JAR 단계를 제출하려면 AWS CLI

클러스터를 시작하고를 사용하여 사용자 지정 JAR 단계를 제출하려면 `--steps` 파라미터와 함께 `create-cluster` 하위 명령을 AWS CLI입력합니다.

- 클러스터를 시작하고 사용자 지정 JAR 단계를 제출하려면 다음 명령을 입력하고 *myKey*를 EC2 키 페어의 이름으로 바꾸며 *amzn-s3-demo-bucket*을 버킷 이름으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbuck
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`--instance-groups` 파라미터를 사용하지 않고 인스턴스 수를 지정하면 단일 프라이머리 노드가 시작되고 나머지 인스턴스는 코어 노드로 시작됩니다. 모든 노드에서 이 명령에 지정한 인스턴스 유형을 사용합니다.

> **Note**
>
> Amazon EMR 서비스 역할과 EC2 인스턴스 프로파일을 아직 생성하지 않았다면 `aws emr create-default-roles` 하위 명령을 입력하기 전에 `create-cluster`를 입력하여 생성합니다.

에서 Amazon EMR 명령을 사용하는 방법에 대한 자세한 내용은 섹션을 AWS CLI참조하세요https://docs.aws.amazon.com/cli/latest/reference/emr.

타사 종속성

때로는 프로그램과 함께 사용하기 위해 MapReduce 클래스 경로 JAR에 포함할 필요가 있을 수 있습니다. 이 작업에는 두 가지 옵션이 있습니다.

- 를 사용하여 클러스터 시작 및 사용자 지정 JAR 단계 제출 AWS CLI의 절차의 경우 `--libjars s3://URI_to_JAR`을 단계 옵션에 포함합니다.
- `mapred-site.xml`에서 수정된 `mapreduce.application.classpath` 설정을 사용하여 클러스터를 시작합니다. `mapred-site` 구성 분류를 사용하십시오. 를 사용하는 단계를 사용하여 클러스터를 생성하려면 다음과 AWS CLI같습니다.

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Hue Name=Hive Name=Pig --use-default-roles \
--instance-type m5.xlarge --instance-count 2  --ec2-attributes KeyName=myKey \
--steps Type=CUSTOM_JAR,Name="Custom JAR
 Step",ActionOnFailure=CONTINUE,Jar=pathtojarfile,Args=["pathtoinputdata","pathtooutputbucket
 \
```

```
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

`myConfig.json`:

```
[
    {
      "Classification": "mapred-site",
      "Properties": {
        "mapreduce.application.classpath": "path1,path2"
      }
    }
  ]
```

쉼표로 구분된 경로 목록을 각 작업의 JVM 클래스 경로에 추가해야 합니다.

복원된 객체 읽기

Amazon EMR 릴리스 7.2.0 이상에서는 S3A 프로토콜을 사용하여 테이블의 S3 위치에서 복원된 Glacier 객체를 읽을 수 있습니다. 이전 릴리스의 엔진은 Glacier와 Glacier Deep Archive 파일을 구분하지 않습니다. 즉, S3A를 사용하여 진행 중인 Glacier 파일에 액세스하려고 하면 `AmazonS3Exception`이 발생합니다. 이 읽기 작업은 아카이브된 Glacier 파일이 여전히 복원 중인 경우 해당 파일을 무시합니다. 이 동작을 활성화하려면 `fs.s3a.glacier.read.restored.objects` 설정을 사용합니다. 이 설정은 세 가지 값일 수 있습니다.

- READ_ALL – 이 값은 Amazon EMR이 Amazon S3에서 검색된 스토리지 클래스를 고려하지 않아야 함을 나타냅니다. 이는 기본 설정 동작입니다.
- SKIP_ALL_GLACIER – 이 값은 Amazon EMR이 Glacier 스토리지 클래스로 태그가 지정된 모든 S3 객체를 무시하고 다른 모든 객체를 검색해야 함을 나타냅니다. 이는 Glacier 객체와 관련된 Amazon Athena의 기본 동작입니다.
- READ_RESTORED_GLACIER_OBJECTS – 이 값은 Amazon EMR이 Glacier 객체의 복원된 상태를 확인해야 함을 나타냅니다. Amazon EMR이 객체를 복원할 수 있는 경우 일반 S3 객체처럼 읽을 수 있습니다. 그렇지 않으면 Amazon EMR은 S3 Glacier의 객체를 무시합니다.

예시

Spark

Spark를 사용하는 경우 복원된 객체를 읽으려면 다음 구성을 사용합니다.

```
--conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

spark-sql을 사용하는 경우 대신 다음 구성을 사용합니다.

```
spark-sql --conf spark.hadoop.fs.s3a.glacier.read.restored.objects=<value>
```

Flink

Flink를 사용하는 경우 `flink-conf.yaml` 파일에서 구성을 설정할 수 있습니다.

```
fs.s3a.glacier.read.restored.objects: <value>
```

`flink-conf` 분류를 설정할 수도 있습니다.

```
[
 {
  "Classification": "flink-conf",
  "Properties": {
  "fs.s3a.glacier.read.restored.objects":"<value>"
  }
 }
]
```

Hive

Hive를 사용하는 경우 `hive-site.xml` 파일에서 구성을 설정합니다.

```
<property>
 <name>fs.s3a.glacier.read.restored.objects</name>
 <value><value></value>
</property>
```

또한 Hive CLI를 사용하여 `--hiveconf` 속성을 설정할 수도 있습니다.

```
hive --hiveconf fs.s3a.glacier.read.restored.objects=<value>
```

고려 사항

S3 Glacier에서 복원된 객체를 읽을 때 다음 고려 사항에 유의합니다.

- S3A 스키마 또는 S3AFileSystem을 사용하여 데이터에 액세스하는 경우에만 복원된 객체를 읽을 수 있습니다.
- 복원된 Glacier 객체를 읽는 경우 Amazon EMR은 객체 자체를 복원하지 않습니다. 이렇게 하려면 AWS CLI 또는 AWS SDK를 사용해야 합니다.

YARN 컨테이너에 대해 비균일 메모리 액세스 인식 활성화

Amazon EMR 버전 6.x 이상에서 클러스터 데이터의 다중 처리에 비균일 메모리 액세스(NUMA)를 사용할 수 있습니다. NUMA는 프로세서가 다른 프로세서에 있는 메모리보다 더 빠르게 자체 로컬 메모리에 액세스하거나 프로세서 간에 공유되는 메모리에 액세스할 수 있는 컴퓨터 메모리 설계 패턴입니다. YARN 컨테이너는 모든 후속 메모리 할당을 지원하는 특정 NUMA 노드에 바인딩할 수 있으므로 NUMA의 성능을 개선합니다. 이를 통해 클러스터가 원격 메모리에 액세스해야 하는 시간이 줄어듭니다.

워커 노드 시스템이 다중 NUMA 노드인 경우 YARN 컨테이너에 대한 NUMA 지원을 켤 수 있습니다. 워커 노드가 단일 NUMA 노드인지, 다중 NUMA 노드인지 확인하려면 다음 명령을 실행합니다.

```
lscpu | grep -i numa
NUMA node(s): 2
```

일반적으로 12x보다 큰 인스턴스에는 두 개의 NUMA 노드가 있습니다. 금속 인스턴스에는 적용되지 않습니다.

YARN 컨테이너에 대한 NUMA 인식을 켜는 방법

1. Amazon EMR 6.x 클러스터에서 다음 yarn-site 구성을 사용합니다.

```
[
{
 "classification":"yarn-site",
  "properties":{
   "yarn.nodemanager.linux-container-executor.nonsecure-mode.local-user":"yarn",
   "yarn.nodemanager.linux-container-executor.group":"yarn",
   "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor",
   "yarn.nodemanager.numa-awareness.enabled":"true",
   "yarn.nodemanager.numa-awareness.numactl.cmd":"/usr/bin/numactl",
   "yarn.nodemanager.numa-awareness.read-topology":"true"
  },
```

```
  "configurations":[]
 }
 ]
```

2. 클러스터에서 다음 부트스트랩 작업을 제공합니다.

```
#!/bin/bash

sudo yum -y install numactl
echo 1 | sudo tee /proc/sys/kernel/numa_balancing

echo "banned.users=mapred,bin,hdfs" >> /etc/hadoop/conf/container-executor.cfg
rm -rf /var/log/hadoop-yarn/
sudo chown -R yarn:hadoop /var/log/hadoop-yarn/
sudo chmod 755 -R /var/log/hadoop-yarn/

sudo chmod 6050 /etc/hadoop/conf/container-executor.cfg

mkdir /mnt/yarn && sudo chmod 755 -R /mnt/yarn && sudo chown -R yarn:hadoop /mnt/
yarn
mkdir /mnt1/yarn && sudo chmod 755 -R /mnt1/yarn && sudo chown -R yarn:hadoop /
mnt1/yarn
mkdir /mnt2/yarn && sudo chmod 755 -R /mnt2/yarn && sudo chown -R yarn:hadoop /
mnt2/yarn
```

3. 모든 컨테이너에서 NUMA를 인식해야 합니다. NUMA 플래그를 사용하여 각 컨테이너의 Java 가상 머신(JVM)에 알릴 수 있습니다. 예를 들어 MapReduce 작업에 NUMA를 사용하도록 JVM에 알리려면 `mapred-site.xml`에서 다음 속성을 추가합니다.

```
<property>
 <name>mapreduce.reduce.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
<property>
 <name>mapreduce.map.java.opts</name>
 <value>-XX:+UseNUMA</value>
</property>
```

4. NUMA가 켜져 있는지 확인하려면 다음 명령을 사용하여 NodeManager 로그 파일을 검색합니다.

```
grep "NUMA resources allocation is enabled," *
```

NodeManager가 컨테이너에 NUMA 노드 리소스를 할당했는지 확인하려면 다음 명령을 사용하여 NodeManager 로그를 검색합니다. 이때 *<container_id>*를 자체 컨테이너 ID로 대체합니다.

```
grep "NUMA node" | grep <container_id>
```

YARN 컨테이너 빈 패킹

Amazon EMR 버전 7.9.0부터 YARN의 다중 노드 배치 정책을 기반으로 구축된 YARN 용량 스케줄러에 컨테이너 빈 패키징 정책을 사용할 수 있습니다. 이 기능은 기본적으로 비활성화되어 있지만 활성화되면 YARN은 구성에서 정의한 사전 정의된 패킹 임계값을 준수하면서 다른 클러스터 노드로 확장하기 전에 단일 노드를 컨테이너로 채우는 것을 우선시합니다`yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage`.

컨테이너 빈 패키징 정책은 기본 균일한 컨테이너 할당 전략에 비해 몇 가지 이점을 제공합니다.

- 클러스터 리소스 조각화를 줄입니다.
- 노드에 사용 가능한 리소스가 있는 경우 제한된 수의 노드에서 컨테이너를 시작하여 다른 노드를 유휴 상태로 두어 클러스터 스케일 다운 작업을 가속화할 수 있으므로 클러스터를 동적으로 확장할 때 비용을 절감할 수 있습니다.

기능 활성화

Amazon EMR에서 컨테이너 빈 패키징 기능을 활성화하려면 다음 YARN 사이트 분류를 추가할 수 있습니다.

```
[
 {
  "Classification": "yarn-site",
  "Properties": {
  "yarn.scheduler.capacity.multi-node-placement.container.bin-packing.percentage": 
 "integer value from 1-100"
  }
 }
]
```

고려 사항

- 이 기능은 YARN 용량 스케줄러에서만 사용할 수 있습니다.
- 이 기능을 활성화하면 YARN 다중 노드 배치 예약 전략이 자동으로 활성화됩니다.
- 제한된 수의 노드에서 리소스 사용률이 집중되어 성능이 저하될 수 있습니다.
- 이 기능을 사용하면 사용자 지정 Auto Scaling 정책이 관리형 조정 정책에 비해 더 나은 축소 작업을 보여줍니다.

Hadoop 버전 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Hadoop의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Hadoop 버전 정보

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-7.9.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.8.0	3.4.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-nameno

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
		de, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.7.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.6.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-7.5.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.4.0	3.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-7.3.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.2.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.36.2	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-7.1.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-7.0.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.15.0	3.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.14.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.13.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.12.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.11.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.11.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.10.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.10.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.9.1	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.9.0	3.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.8.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.8.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.7.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.36.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.36.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.6.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.35.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.5.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.4.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.3.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.3.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.2.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.2.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.1.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.1.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-6.0.1	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-6.0.0	3.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.34.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.33.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.33.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.32.1	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.32.0	2.10.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.31.1	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.31.0	2.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.30.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.30.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.29.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.28.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.28.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.27.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.27.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.26.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.25.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.24.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.24.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.23.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.23.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.22.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.21.2	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.21.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.21.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.20.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.20.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.19.1	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.19.0	2.8.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.18.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.18.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.2	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.17.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.17.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.16.1	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.16.0	2.8.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.15.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.15.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.14.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.14.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.14.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.13.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.13.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.3	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.12.2	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.12.1	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.12.0	2.8.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.11.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.11.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.11.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.10.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.10.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.9.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.8.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.8.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.7.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.7.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.6.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.6.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server
emr-5.5.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.5.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.5.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.5.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.4.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.3.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.3.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.2.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.2.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.2.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.1.0	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.0.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-5.0.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-5.0.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.6	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.9.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.9.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.9.1	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.5	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.4	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.8.3	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.2	2.7.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.8.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.8.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.4	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.3	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.7.2	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.7.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.6.1	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.6.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.5.0	2.7.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.4.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.3.0	2.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.2.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

Amazon EMR 릴리스 레이블	하둡 버전	Hadoop과 함께 설치된 구성 요소
emr-4.1.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager
emr-4.0.0	2.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager

버전별 Hadoop 릴리스 정보

주제

- Amazon EMR 7.9.0 - 하둡 릴리스 정보
- Amazon EMR 7.8.0 - 하둡 릴리스 정보
- Amazon EMR 7.7.0 - 하둡 릴리스 정보
- Amazon EMR 7.6.0 - 하둡 릴리스 정보
- Amazon EMR 7.5.0 - 하둡 릴리스 정보
- Amazon EMR 7.4.0 - 하둡 릴리스 정보
- Amazon EMR 7.3.0 - Hadoop 릴리스 정보
- Amazon EMR 7.2.0 - Hadoop 릴리스 정보

- Amazon EMR 6.6.0 - Hadoop 릴리스 정보

Amazon EMR 7.9.0 - 하둡 릴리스 정보

Amazon EMR 7.9.0 - 하둡 변경 사항

유형	설명
새로운 기능	자동 구성 매핑: EMRFS 파일 시스템 구성은 해당 S3A 파일 시스템 구성에 자동으로 적용되므로 EMRFS에서 S3A로 원활하게 마이그레이션할 수 있습니다.
새로운 기능	YARN Conatainer bin-packing: 애그리시브 축소를 위한 예약 정책
백포트	YARN-11752: Global Scheduler: 컨테이너 할당 시간 개선
버그 수정	S3A: 만료된 자격 증명을 반환하는 IAMCredentialsProvider 수정
버그 수정	전송 중 데이터 암호화가 활성화된 클러스터의 `yarn.log.server.url` 구성 값 수정

Amazon EMR 7.8.0 - 하둡 릴리스 정보

Amazon EMR 7.8.0 - 하둡 변경 사항

유형	설명
업그레이드	하둡이 버전 3.4.1로 업그레이드되었습니다.

Amazon EMR 7.7.0 - 하둡 릴리스 정보

Amazon EMR 7.7.0 - 하둡 변경 사항

유형	설명
새로운 기능	S3 접두사 목록을 사용하여 S3A GlobStatus 호출 최적화 S3
백포트	YARN-7327: 용량 스케줄러에 대해 기본적으로 비동기 예약 활성화
백포트	YARN-10058: 스케줄러 중단 방지를 위해 비동기 예약 스레드에 대해 알 수 없는 예외 처리
백포트	YARN-11732: CapacityScheduler에 대해 SchedulerNode#reservedContainer를 호출할 때 잠재적 NPE 수정
백포트	YARN-11560: 일정으로 다중 노드가 비동기적으로 활성화된 경우 NPE 버그 수정
백포트	YARN-11191: GlobalScheduler refreshQueues의 강력한 교착 상태 수정
백포트	YARN-11041: queuePath의 모든 발생을 새 QueuePath 클래스로 대체
백포트	YARN-11660: SingleConstraintAppPlacementAllocator의 성능 회귀 수정
백포트	HADOOP-19116: CVE-2024-23944로 인한 zookeeper 클라이언트 3.8.4 업데이트.

유형	설명
백포트	HADOOP-19115: CVE-2023-52428로 인해 nimbus-jose-jwt 9.37.2로 업그레이드합니다. CVE-2023-52428
백포트	HADOOP-19024: bouncycastle jdk18 1.77 사용
백포트	HADOOP-19123: CVE로 인한 commons-c onfiguration2 2.10.1 업데이트
백포트	HADOOP-19114: CVEs로 인해 commons-c ompress 1.26.1로 업그레이드
백포트	HADOOP-19237: CVEs로 인해 dnsjava 3.6.1로 업그레이드
새로운 기능	S3A에 S3 요청 감사 추가
백포트	HADOOP-17609: OpenSSL 네이티브 코드에 대해 SM4 지원을 선택 사항으로 설정
백포트	HADOOP-18583: hadoop 체크네이티브가 openssl 3.x 로드 실패
새로운 기능	S3A 역할 매핑에 대한 지원 추가

Amazon EMR 7.7.0 - 하둡 기능

- 비동기식 컨테이너 예약은 용량 스케줄러의 기본 예약 전략으로, 컨테이너 할당 속도를 최적화하도록 설계되었습니다.
- S3A 파일 시스템은 S3 접두사 목록을 사용하여 목록 작업을 가속화하는 glob 상태 호출에 대한 최적화를 도입합니다. 기본적으로이 기능은 비활성화되어 있으며 core-site.xml 파

일`fs.s3a.prefix.listing.in.glob.status.enabled=true`에서를 구성하여 활성화 할 수 있습니다. 이 최적화를 활성화하면와 같은 globstatus 호출에 대한 서버 측 필터링이 가능`fs.globstatus("s3://`*`bucket`*`/a*")`하므로 로 시작하는 객체만 나열하여 목록 성능이 향상됩니다`"a"`.

- 활성화된 경우 S3 요청 감사를 S3A에 추가fileSystemOwner 객체의 정보를 사용하여 userAgent 문자열을 S3 요청을 수행하는 사용자 및 사용자 그룹 필드로 채웁니다.
- S3A는 사용자, 그룹 또는 S3 접두사를 기반으로 사용할 IAM 역할을 결정하는 데 도움이 되는 역할 매핑에 대한 지원을 추가합니다.

Amazon EMR 7.6.0 - 하둡 릴리스 정보

Amazon EMR 7.6.0 - 하둡 변경 사항

유형	설명
새로운 기능	S3A 자격 증명 벤딩 프레임워크
새로운 기능	Hive에 대한 S3A 매직 커미터 지원
백포트	HADOOP-19309: S3A: 소스 파일에 파일 체계가 포함되지 않은 경우 CopyFromLocalFile 작업이 실패합니다.
백포트	HADOOP-19050 S3A:S3 액세스 권한 부여 지원
백포트	HADOOP-18708 S3A: S3 클라이언트 측 암호화(CSE) 지원

Amazon EMR 7.6.0 - 하둡 기능

- `yarn.scheduler.increment-allocation-mb` 및에 대한 기본 구성 설정이 32에서 1로 수정`yarn.scheduler.minimum-allocation-mb`되어 YARN에서 컨테이너 리소스 정규화가 효과적으로 비활성화됩니다.

- 요청 수준 자격 증명 벤딩을 지원하도록 하둡의 S3A 파일 시스템이 개선되었습니다. 이 기능을 사용하면 모든 Amazon S3 요청은를 활용합니다`AwsCredentialsProvider`. 이렇게 하면 각 요청에 대한 자격 증명을 동적으로 제공하여 유연성과 보안이 향상됩니다.
- 하둡의 S3A 파일 시스템은 S3 클라이언트 측 암호화를 지원합니다.

Amazon EMR 7.5.0 - 하둡 릴리스 정보

Amazon EMR 7.5.0 - 하둡 변경 사항

유형	설명
버그 수정	fs.file.impl을 빈 값으로 주석 처리했습니다.
백포트	HADOOP-19286: S3 리전/엔드포인트가 설정된 경우 S3A 리전 간 액세스 지원 S3
개선	EMR-EC2의 S3A 커넥터에 대한 자동 S3 리전 구성 설정
개선	S3A에서 HeadObject 호출 수 줄이기

Amazon EMR 7.5 릴리스에서 Spark의 S3A 커넥터는 3TB TPC-DS parquet 데이터 세트를 사용한 벤치마크에서 알 수 있듯이 EMRFS와 유사한 읽기 성능을 보여줍니다.

Amazon EMR 7.5.0 - 하둡 기능

- S3 리전 구성`fs.s3a.endpoint.region`은 EMR-EC2 배포용 S3A 커넥터를 사용하여 EMR 클러스터가 시작되는 리전으로 자동 설정됩니다. EMR-EC2
- Amazon S3 교차 버킷 리전 액세스는 S3A 커넥터에 대해 기본적으로 활성화됩니다. 를 설정하여 수정할 수 있습니다`fs.s3a.cross.region.access.enabled=`*`true or false`*.

Amazon EMR 7.4.0 - 하둡 릴리스 정보

Amazon EMR 7.4.0 - 하둡 변경 사항

유형	설명
업그레이드	하둡 버전이 3.4.0으로 업그레이드되었습니다. OSS 릴리스 정보를 참조하세요.
버그 수정	FairScheduler에 대한 음의 보류 중 및 할당된 Yarn 지표 수정
버그 수정	YARN-11702: 컨테이너 할당에 대한 Yarn 수정
버그 수정	노드 축소 시 레이스 조건 처리 개선
개선	HADOOP-18679: 파일의 대량/페이지 삭제를 위한 API 추가
개선	HADOOP-19203: IOEs UncheckedIOExceptions로 높이기 위한 WrappedIO BulkDelete API
개선	HADOOP-19205: S3A: v1 SDK보다 초기화/닫기가 느림
개선	HADOOP-19161: S3A: 성능 플래그 목록을 가져오기 위한 fs.s3a.performance.flags 옵션
개선	HADOOP-19072: S3A: fs.s3a.performance.flags for mkdir를 사용하여 스토어에 대한 최적화 확장

Amazon EMR 7.4.0 - 하둡 기능

Amazon EMR 7.4.0의 새로운 하둡 기능은 다음 목록을 참조하세요.

- 기본 구성 값은 최적의 성능을 위해 미세 조정되었습니다.
  - mapreduce.input.fileinputformat.list-status.num-threads=10 – 1에서 증가했습니다.
  - fs.s3a.block.size=64M – 32M에서 증가했습니다.
  - fs.s3a.multipart.size=128M – 64M에서 증가했습니다.
- S3A 파일 시스템을 사용하여 MapReduce 작업을 가속화하기 위한 기본 제공 Out-of-the-box 성능 향상 최적화.

Amazon EMR 7.3.0 - Hadoop 릴리스 정보

Amazon EMR 7.3.0 - Hadoop 변경 사항

유형	설명
새로운 기능	HADOOP-19197: S3A 지원 AWS KMS 암호화 컨텍스트
새로운 기능	HADOOP-18980: S3A 자격 증명 제공업체 재매핑: 확장 가능
개선	HADOOP-18808: debug() 메서드를 추가할 LogExactlyOnce

Amazon EMR 7.3.0 - Hadoop 기능

Amazon EMR 7.3.0의 새로운 Hadoop 기능은 다음 목록을 참조하세요.

- 기본적으로 보안 구성을 통해 전송 중 데이터 암호화가 활성화된 클러스터는 Hadoop YARN(Resource Manager, Node Manager, Timeline Server, WebApplicationProxy), HDFS(NameNode, DataNode, Journal Node, DFSZKFailoverController), Map Reduce(JobHistory Server 및 MapReduce Shuffle 포트)에 대해 TLS 1.3을 실행합니다.

Amazon EMR 7.2.0 - Hadoop 릴리스 정보

Amazon EMR 7.2.0 - Hadoop 변경 사항

유형	설명
새로운 기능	HADOOP-18850: S3A: AWS KMS 키를 사용한 이중 계층 서버 측 암호화 활성화
새로운 기능	S3 Glacier 복원된 객체 읽기 지원
개선	Spark 작업 실행 시 코어를 기반으로 S3A 최대 연결 동적 조정
개선	HADOOP-18797: S3A Magic Committer를 사용한 동시 쓰기 지원
개선	HADOOP-19047: S3A Magic Commits의 InMemory 추적 지원
개선	MagicV2 Committer에 'Export', 'Publish' 및 AbortPublish' 메서드 추가
개선	S3A에서 ReadFullyIntoBuffers FileSystem API 구현
버그 수정	HADOOP-18793: S3A StagingCommitter가 staging-uploads 디렉터리를 정리하지 않음
업그레이드	HADOOP-19008: S3A: aws-sdk 버전을 2.21.41로 업데이트
업그레이드	

유형	설명
	HADOOP-18613: ZooKeeper를 버전 3.8.3으로 업그레이드

Amazon EMR 7.2.0 - Hadoop 기능

Amazon EMR 7.2.0의 새로운 Hadoop 기능은 다음 목록을 참조하세요.

- 복원된 객체를 읽기 위한 S3 glacier 지원 - Amazon EMR 릴리스 7.2.0 이상에서는 S3A 프로토콜을 사용하여 테이블의 S3 위치에서 복원된 Glacier 객체를 읽을 수 있습니다. 이 읽기 작업은 아카이브된 Glacier 파일이 여전히 복원 중인 경우 해당 파일을 무시합니다.
- S3A Magic Committer를 사용한 동시 쓰기 지원 - Amazon EMR 7.2.0에서는 S3A Magic Committer를 사용한 동시 쓰기 지원을 도입합니다. 이 기능은 HADOOP-18797의 일부로 OSS Hadoop 커뮤니티에 추가되었습니다.

Amazon EMR 6.6.0 - Hadoop 릴리스 정보

Amazon EMR 6.6.0 - Hadoop 변경

유형	설명
버그	BZip2 텍스트 파일을 읽을 때 중복되는 레코드를 수정했습니다.
백포트	HADOOP-18136: 누락된.tar 파일의 `FileUtils.unTar()` 처리 확인
백포트	HADOOP-17627: branch-3.2 HADOOP-17371, HADOOP-17621, HADOOP-17625로 백포트하여 Jetty를 9.4.39로 업데이트
백포트	HADOOP-17655: Jetty를 9.4.40으로 업그레이드
백포트	HADOOP-17796: jetty 버전을 9.4.43으로 업그레이드

유형	설명
백포트	HADOOP-17661: mvn versions:set에서 pom.xml 구문 분석 실패
백포트	HADOOP-17236: snakeyaml을 1.26로 업그레이드하여 CVE-2017-18640 완화
백포트	HADOOP-16717: Log4jLoggerAdapter에서 GenericsUtil isLog4jLogger 종속성 제거
백포트	HADOOP-17633: CVE 때문에 json-smart를 2.4.2로, nimbus-jose-jwt를 9.8로 전환
백포트	HADOOP-17844: JSON smart를 2.4.7로 업그레이드
백포트	HADOOP-17972: branch-3.2에서 HADOOP-17683 백포트(commons-io를 2.8.0으로 업데이트)
백포트	HADOOP-16555: commons-compress를 1.19로 업데이트
백포트	HADOOP-17370: commons-compress를 1.21로 업그레이드
백포트	HADOOP-17096: ZStandardCompressor 입력 버퍼 오프셋 수정
백포트	HADOOP-17112: 커미터를 통해 s3a에 파일을 저장할 때 경로에 공백이 허용되지 않음
백포트	HADOOP-13500: 구성 속성 객체의 반복 동기화
백포트	HDFS-14099: ZStandardDecompressor에서 여러 프레임을 압축 해제할 때 알 수 없는 프레임 디스크립터
백포트	HDFS-16410: OfflineEditsXmlLoader에서 안전하지 않은 Xml 구문 분석

유형	설명
백포트	HDFS-14498: LeaseManager는 생성이 실패한 파일에서 영구적으로 반복될 수 있음
백포트	HDFS-15290: NameNode 시작 중 HttpServer에서 NPE
백포트	HDFS-15293: 체크포인트 수신 시 fsimage 수락 조건 완화
백포트	HDFS-12979: StandbyNode에서 검사 이후 ObserverNode에 FsImage를 업로드해야 함
백포트	YARN-10538: AM에 반환되는 업데이트된 노드 목록에 재지정 노드 추가
백포트	YARN-10472: YARN-10314(YarnClient에서 음영 표시된 클라이언트 jar에서만 WebSocket Exception에 대해 NoClassDefFoundError 생성)를 branch-3.2로 백포트
백포트	YARN-9968: `NullPointerException` 때문에 `NodeManager` 에서 퍼블릭 로컬라이저가 종료됨
백포트	YARN-10651: `CapacityScheduler` 가 `AbstractYarnScheduler.updateNodeResource()` 의 NPE와 충돌함
백포트	YARN-9339: 앱을 새 대기열로 이동한 후 앱에서 보류 중인 지표가 올바르지 않음
백포트	YARN-10438: `ClientRMService#getContainerReport()` 에서 null `containerId` 처리
백포트	YARN-7266: `RollingLevelDb` 파일이 손상되거나 누락된 경우 ATS 1.5가 시작되지 않음

유형	설명
백포트	YARN-9063: `RollingLevelDb` 파일이 손상되거나 누락된 경우 ATS 1.5가 시작되지 않음
백포트	YARN-9848: YARN-4946을 되돌립니다(RM은 로그 집계가 터미널 상태가 아닌 경우 애플리케이션을 `COMPLETED` 로 간주할 수 없음).

Apache HBase

HBase는 Apache Software Foundation의 하둡 프로젝트의 일부로 개발된 오픈 소스 비관계형 분산 데이터베이스입니다. HBase는 하둡 분산 파일 시스템(HDFS)을 기반으로 실행되어 하둡 에코시스템을 위한 비관계형 데이터베이스 기능을 제공합니다. HBase는 Amazon EMR 릴리스 버전 4.6.0 이상에 포함되어 있습니다.

HBase는 하둡과 원활하게 작동하여 파일 시스템을 공유하고 MapReduce 프레임워크 및 실행 엔진에 직접 입력 및 출력할 수 있습니다. HBase는 또한 Apache Hive와 통합하여 HBase 테이블을 통해 SQL과 같은 쿼리를 작성하고 Hive 기반 테이블 및 Java Database Connectivity(JDBC)에 대한 지원과 결합할 수 있습니다. HBase에 대한 자세한 내용은 Apache 웹 사이트의 Apache HBase 및 HBase 설명서를 참조하십시오. HBase를 Hive와 함께 사용하는 방법의 예는 AWS 빅 데이터 블로그 게시물 NoSQL 결합과 Amazon EMR에서 Apache HBase 및 Apache Hive를 사용한 대량 병렬 분석을 참조하세요.

Amazon EMR 기반 HBase를 사용하면 HBase 데이터를 Amazon Simple Storage Service(S3)에 직접 백업하고 HBase 클러스터를 시작할 때 이전에 생성한 백업에서 복원할 수도 있습니다. Amazon EMR은 데이터 지속성 및 재해 복구를 위해 Amazon S3와 통합할 수 있는 추가 옵션을 제공합니다.

- Amazon S3 기반 HBase - Amazon EMR 버전 5.2.0 이상에서는 Amazon S3 기반 HBase를 사용하여 클러스터의 HBase 루트 디렉터리와 메타데이터를 Amazon S3에 직접 저장할 수 있습니다. 이후에는 Amazon S3에서 루트 디렉터리 위치를 가리켜서 새 클러스터를 시작할 수 있습니다. 읽기 전용 복제본 클러스터를 제외하고 한 번에 한 클러스터에서만 Amazon S3의 HBase 위치를 사용할 수 있습니다. 자세한 내용은 Amazon S3 기반(Amazon S3 스토리지 모드) 단원을 참조하십시오.
- HBase 읽기 전용 복제본 - Amazon S3 기반 HBase가 있는 Amazon EMR 버전 5.7.0 이상에서는 읽기 전용 복제본 클러스터를 지원합니다. 읽기 전용 복제본 클러스터에서는 읽기 전용 작업을 위해 기본 클러스터의 스토어 파일 및 메타데이터에 읽기 전용 권한으로 액세스할 수 있습니다. 자세한 내용은 읽기 전용 복제본 클러스터 사용 단원을 참조하십시오.
- HBase 스냅샷 - Amazon S3 기반 HBase 대신, EMR 버전 4.0 이상을 사용하여 HBase 데이터에 대한 스냅샷을 Amazon S3에 직접 생성한 다음, 스냅샷을 사용하여 데이터를 복구할 수 있습니다. 자세한 내용은 HBase 스냅샷 사용 단원을 참조하십시오.

> **Important**
>
> Amazon EMR HBase 클러스터 조정의 경우 HBase 클러스터의 Managed Scaling 또는 사용자 지정 정책을 통한 조정을 사용하지 않는 것이 좋습니다.

다음 테이블에는 Amazon EMR이 HBase를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 HBase의 버전이 나열되어 있습니다.

이 릴리스에서 HBase와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 HBase 버전 정보

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.9.0	HBase 2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

다음 테이블에는 Amazon EMR이 HBase를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 HBase의 버전이 나열되어 있습니다.

이 릴리스에서 HBase와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 HBase 버전 정보

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.15.0	HBase 2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

> **Note**
>
> Apache HBase HBCK2는 HBase 리전 및 시스템 테이블을 복구하기 위한 별도의 운영 도구입니다. Amazon EMR 버전 6.1.0 이상에서는 프라이머리 노드의 `/usr/lib/hbase-operator-tools/`에 hbase-hbck2.jar가 제공됩니다. 도구를 빌드하고 사용하는 방법에 대한 자세한 내용은 HBase HBCK2를 참조하십시오.

다음 테이블에는 Amazon EMR이 HBase를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 HBase의 버전이 나열되어 있습니다.

이 릴리스에서 HBase와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 HBase 버전 정보

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.36.2	HBase 1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

주제

- HBase를 포함하는 클러스터 생성
- Amazon S3 기반(Amazon S3 스토리지 모드)
- Amazon EMR에 대한 미리 쓰기 로그(WAL)
- HBase 쉘 사용
- Hive를 사용하여 HBase 테이블 액세스
- HBase 스냅샷 사용
- HBase 구성
- HBase 사용자 인터페이스 보기
- HBase 로그 파일 보기
- Ganglia를 사용하여 HBase 모니터링
- Amazon CloudWatch를 사용하여 EMR HBase 모니터링
- 이전 HBase 버전에서 마이그레이션

- [HBase 릴리스 기록]

HBase를 포함하는 클러스터 생성

이 섹션의 절차에서는 AWS Management Console 및를 사용하여 클러스터를 시작하는 기본 사항을 다룹니다 AWS CLI. Amazon EMR 클러스터를 계획, 구성 및 시작하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서의 [클러스터 계획 및 구성]을 참조하세요.

콘솔을 사용하여 HBase를 포함하는 클러스터 생성

콘솔을 사용하여 클러스터를 시작하는 빠른 단계는 Amazon EMR 관리 안내서에서 [Amazon EMR 시작하기]를 참조하세요.

콘솔을 사용하여 설치된 HBase와 함께 클러스터를 시작하려면

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. 소프트웨어 구성에서 Amazon 릴리스 버전 4.6.0 이상을 선택합니다(최신 버전 권장). HBase 및 기타 원하는 애플리케이션을 선택합니다.
4. Amazon EMR 버전 5.2.0 이상의 경우 HBase 스토리지 설정 아래에서 HDFS 또는 S3를 선택합니다. 자세한 내용은 [Amazon S3 기반(Amazon S3 스토리지 모드)] 단원을 참조하십시오.
5. 필요에 따라 다른 옵션을 선택한 다음 클러스터 생성을 선택합니다.

를 사용하여 HBase로 클러스터 생성 AWS CLI

다음 명령을 사용하여 HBase가 설치된 상태로 클러스터를 생성합니다.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

Amazon S3 기반 HBase를 사용할 경우 JSON 구성 객체에 대한 참조와 함께 `--configurations` 옵션을 지정합니다. 구성 객체에는 `hbase.rootdir` 속성을 사용하여 HBase 데이터가 저장된 Amazon S3에서의 위치를 지정하는 `hbase-site` 분류가 포함되어야 합니다. 또한 `hbase` 속성을 사용하여 `s3`을 지정하는 `hbase.emr.storageMode`분류도 포함되어야 합니다. 다음 예제에서는 이러한 구성 설정이 포함된 JSON 코드 조각을 보여줍니다.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3"
        }
    }
]
```

Amazon S3 기반 HBase에 대한 자세한 내용은 Amazon S3 기반(Amazon S3 스토리지 모드) 섹션을 참조하세요. 분류에 대한 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

Amazon S3 기반(Amazon S3 스토리지 모드)

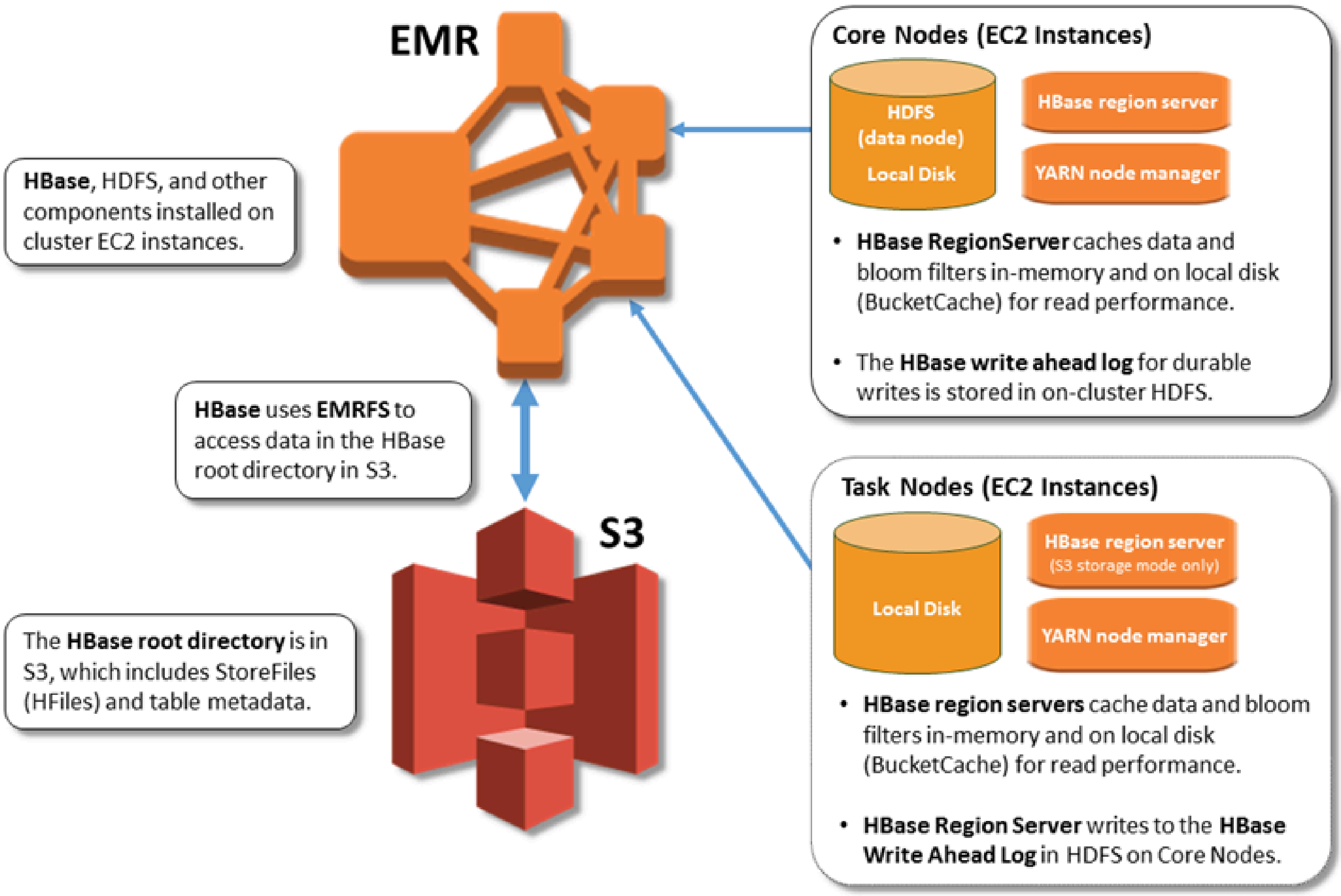
HBase를 Amazon EMR 버전 5.2.0 이상에서 실행할 경우 다음 이점을 제공하는 Amazon S3 기반 HBase를 활성화할 수 있습니다.

- HBase 루트 디렉터리는 HBase 스토어 파일 및 테이블 메타데이터를 포함하여 Amazon S3에 저장됩니다. 이 데이터는 클러스터 외부에서 지속되며 Amazon EC2 가용 영역에서 사용할 수 있으므로 스냅샷이나 다른 방법을 사용하여 복구할 필요가 없습니다.
- Amazon S3의 스토어 파일을 사용하면 데이터 요구 사항 대신 계산 요구 사항에 맞게 Amazon EMR 클러스터 크기를 조정하고 HDFS에서 3배로 복제할 수 있습니다.
- Amazon EMR 버전 5.7.0 이상을 사용하여 읽기 전용 복제본 클러스터를 설정할 수 있습니다. 그러면 Amazon S3에서 데이터의 읽기 전용 복제본을 유지할 수 있습니다. 읽기 전용 복제본 클러스터에서 데이터에 액세스하여 읽기 작업을 동시에 수행할 수 있습니다. 이는 기본 클러스터를 사용할 수 없는 경우에도 마찬가지입니다.

- Amazon EMR 버전 6.2.0~7.3.0에서 영구 HFile 추적은 라는 HBase 시스템 테이블`hbase:storefile`을 사용하여 읽기 작업에 사용되는 HFile 경로를 직접 추적합니다. 이 기능은 기본적으로 활성화되어 있으며, 수동 마이그레이션을 수행할 필요가 없습니다. 7.3.0보다 높은 버전에서는 파일 트래커를 사용하여 HFile 경로를 추적하고 HFile 경로를 저장소 디렉터리 내의 메타 파일에 직접 저장합니다.

> **Note**
>
> 7.4.0 이전 버전의 Amazon EMR을 사용하고 EMR-7.4.0 이상으로 마이그레이션하는 사용자는 이전 HBase 버전에서 마이그레이션을 참조하고 사용 가능한 업그레이드 설명서를 따라 원활한 전환을 보장합니다.

다음 그림은 Amazon S3 기반 HBase에 관련된 HBase 구성 요소를 보여줍니다.



Amazon S3 기반 HBase 활성화

Amazon EMR 콘솔, 또는 Amazon EMR API를 사용하여 Amazon S3에서 HBase를 활성화할 수 있습니다. AWS CLI구성은 클러스터 생성시 옵션입니다. 콘솔을 사용할 경우 고급 옵션을 사용하여 설정을 선택합니다. AWS CLI를 사용할 경우 `--configurations` 옵션을 사용하여 JSON 구성 객체를 제공하십시오. 구성 객체의 속성에서 스토리지 모드와 Amazon S3의 루트 디렉터리 위치를 지정합니다. 지정된 Amazon S3 위치는 Amazon EMR 클러스터와 동일한 리전에 있어야 합니다. 한 번에 하나의 활성 클러스터만 Amazon S3에서 동일한 HBase 루트 디렉터리를 사용해야 합니다. 를 사용한 콘솔 단계 및 자세한 create-cluster 예제는 섹션을 AWS CLI참조하세요[HBase를 포함하는 클러스터 생성](). 구성 객체 예제가 다음 JSON 코드 조각에 표시됩니다.

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir"}
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3"
  }
}
```

> **Note**
>
> Amazon S3 버킷을 HBase의 `rootdir`로 사용하는 경우 Amazon S3 URI의 끝에 슬래시를 추가해야 합니다. 예를 들어, 문제를 방지하기 위해 `"hbase.rootdir: s3://amzn-s3-demo-bucket"` 대신 `"hbase.rootdir: s3://amzn-s3-demo-bucket/"`을 사용해야 합니다.

읽기 전용 복제본 클러스터 사용

Amazon S3 기반 HBase를 사용하여 기본 클러스터를 설정한 후 기본 클러스터와 동일한 데이터에 대한 읽기 전용 액세스 권한을 제공하는 읽기 전용 복제본 클러스터를 생성하여 구성할 수 있습니다. 이는 쿼리 데이터에 동시에 액세스해야 하거나 기본 클러스터를 사용할 수 없는 경우에도 중단 없이 액세

스해야 하는 경우에 유용합니다. 읽기 전용 복제본 기능은 Amazon EMR 버전 5.7.0 이상에서 사용할 수 있습니다.

기본 클러스터와 읽기 전용 복제본 클러스터는 동일한 방법으로 설정하지만 한 가지 중요한 차이점이 있습니다. 두 클러스터는 모두 동일한 `hbase.rootdir` 위치를 가리킵니다. 하지만 읽기 전용 복제본 클러스터의 `hbase` 분류에는 `"hbase.emr.readreplica.enabled":"true"` 속성이 포함되어 있습니다.

읽기 전용 복제본 클러스터는 읽기 전용 작업용으로 설계되었으며 수동 압축 또는 쓰기 작업을 수행할 필요가 없습니다. Amazon EMR 버전 7.4.0 이전의 경우 읽기 전용 복제본 기능을 활성화할 때 읽기 전용 복제본 클러스터에서 압축을 비활성화하는 것이 좋습니다. 기본 클러스터에서 영구 HFile 추적 기능을 활성화하면 읽기 전용 복제본 클러스터가 시스템 테이블을 압축하여 기본 클러스터에 FileNotFoundException이 발생할 수 있으므로 이러한 예방 조치가 필요합니다. 읽기 전용 복제본 클러스터에서 압축을 비활성화하면 기본 클러스터와 읽기 전용 복제본 클러스터 간의 데이터 불일치가 방지됩니다.

예를 들어 주제의 앞부분에서 볼 수 있듯이 기본 클러스터의 JSON 분류를 고려할 때 7.4.0 이전의 EMR 버전에 대한 읽기 전용 복제본 클러스터의 구성은 다음과 같습니다.

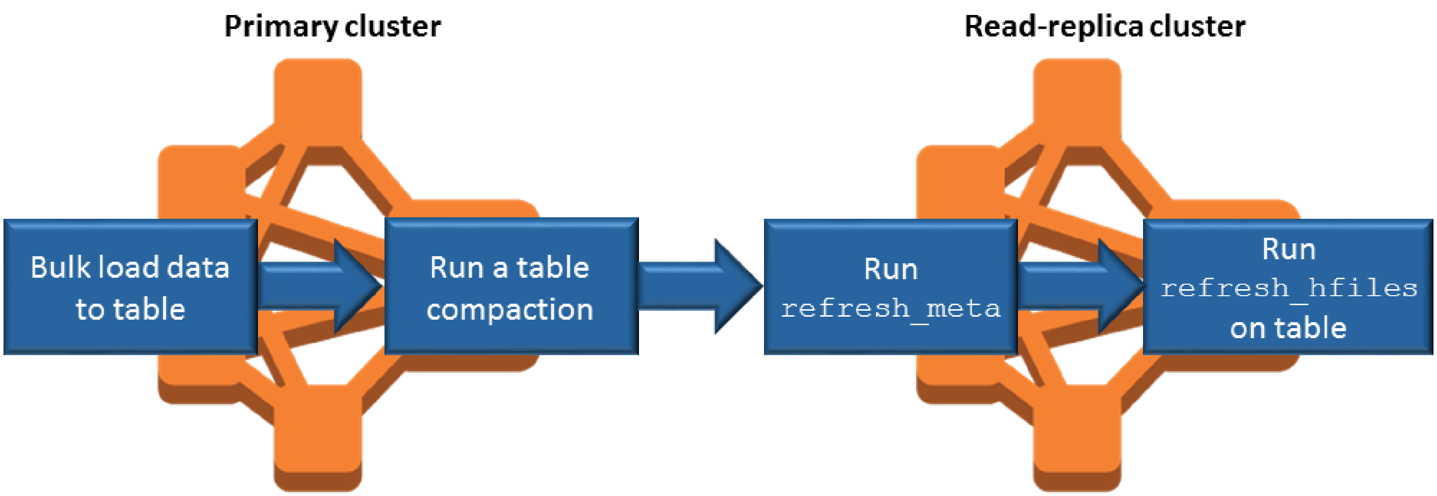
```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.rootdir": "s3://amzn-s3-demo-bucket/my-hbase-rootdir",
    "hbase.regionserver.compaction.enabled": "false"
  }
},
{
  "Classification": "hbase",
  "Properties": {
  "hbase.emr.storageMode":"s3",
  "hbase.emr.readreplica.enabled":"true"
  }
}
```

Amazon EMR 버전 7.3.0 이상에서는 이제 파일 추적 저장 기능을 사용하므로 압축을 비활성화할 필요가 없습니다.

데이터를 추가할 때 읽기 전용 복제본 동기화

읽기 전용 복제본에서는 기본 클러스터에서 Amazon S3에 쓰는 HBase StoreFiles 및 메타데이터를 사용하므로, 읽기 전용 복제본은 Amazon S3 데이터 스토어의 최신 정보만 반영합니다. 다음 지침을 따르면 데이터를 쓸 때 읽기 전용 복제본과 기본 클러스터 사이의 지연 시간을 최소화할 수 있습니다.

- 가능하면 언제든 기본 클러스터에 데이터를 대량 로드합니다. 자세한 내용은 Apache HBase 설명서에서 Bulk loading을 참조하세요.
- 스토어 파일을 Amazon S3에 쓰는 플러시는 데이터를 추가한 후 가능한 빨리 발생해야 합니다. 수동으로 플러시하거나 플러시 설정을 조정하여 지연 시간을 최소화하십시오.
- 압축이 자동으로 실행될 수 있는 경우 수동 압축을 실행하여 압축이 트리거될 때 불일치가 발생하지 않도록 합니다.
- 읽기 전용 복제본 클러스터에서 메타데이터가 변경된 경우(예: HBase 리전 분할 또는 압축을 수행한 경우, 테이블을 추가하거나 제거한 경우) `refresh_meta` 명령을 실행합니다.
- 읽기 전용 복제본 클러스터에서 테이블에서 레코드를 추가하거나 변경할 때 `refresh_hfiles` 명령을 실행합니다.



영구 HFile 추적

영구 HFile 추적은 `hbase:storefile`이라고 하는 HBase 시스템 테이블을 사용하여 읽기 작업에 사용되는 HFile 경로를 직접 추적합니다. HBase에 추가 데이터가 추가되면 새 HFile 경로가 테이블에 추가됩니다. 이렇게 하면 중요한 쓰기 경로 HBase 작업에서 커밋 메커니즘으로 이름 바꾸기 작업이 제거되고, 파일 시스템 디렉터리 나열 대신 `hbase:storefile` 시스템 테이블에서 읽어 HBase 영역을 열 때 복구 시간이 개선됩니다. 이 기능은 Amazon EMR 버전 6.2.0~7.3.0에서 기본적으로 활성화되며 수동 마이그레이션 단계가 필요하지 않습니다.

> **Note**
>
> HBase 스토어파일 시스템 테이블을 사용한 지속적 HFile 추적은 HBase 리전 복제 기능을 지원하지 않습니다. HBase 리전 복제에 대한 자세한 내용은 타임라인이 일치하는 고가용성 읽기를 참조하세요.

영구 HFile 추적 비활성화

영구 HFile 추적은 Amazon EMR 릴리스 6.2.0부터 기본적으로 활성화됩니다. 영구 HFile 추적을 비활성화하려면 클러스터를 시작할 때 다음 구성 재정의를 지정합니다.

```
{
  "Classification": "hbase-site",
  "Properties": {
    "hbase.storefile.tracking.persist.enabled":"false",

 "hbase.hstore.engine.class":"org.apache.hadoop.hbase.regionserver.DefaultStoreEngine"
  }
}
```

> **Note**
>
> Amazon EMR 클러스터를 재구성할 때 모든 인스턴스 그룹을 업데이트해야 합니다.

Storefile 테이블 수동 동기화

새 HFile 인스턴스가 생성되면 storefile 테이블이 최신 상태로 유지됩니다. 그러나 어떤 이유로든 storefile 테이블이 데이터 파일과 동기화되지 않는 경우 다음 명령을 사용하여 데이터를 수동으로 동기화할 수 있습니다.

온라인 리전에서 storefile 테이블 동기화:

```
hbase org.apache.hadoop.hbase.client.example.RefreshHFilesClient <table>
```

오프라인 리전의 storefile 테이블 동기화:

- storefile 테이블 znode를 제거합니다.

```
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
[<tableName>, hbase:namespace]
# The TableName exists in the list
echo "delete /hbase/storefile/loaded/<tableName>" | sudo -u hbase hbase zkcli
# Delete the Table ZNode
echo "ls /hbase/storefile/loaded" | sudo -u hbase hbase zkcli
[hbase:namespace]
```

- 리전을 할당합니다('hbase 쉘'에서 실행).

```
hbase cli> assign '<region name>'
```

- 할당이 실패한 경우.

```
hbase cli> disable '<table name>'
hbase cli> enable '<table name>'
```

storefile 테이블 조정

storefile 테이블은 기본적으로 네 개의 리전으로 분할됩니다. storefile 테이블에 여전히 쓰기 로드가 많은 경우 테이블을 수동으로 더 분할할 수 있습니다.

특정 핫 리전을 분할하려면 다음 명령('hbase shell'에서 실행)을 사용합니다.

```
hbase cli> split '<region name>'
```

테이블을 분할하려면 다음 명령('hbase shell'에서 실행)을 사용합니다.

```
hbase cli> split 'hbase:storefile'
```

파일 추적 저장

기본적으로 FileBasedStoreFileTracker 구현을 사용합니다. 이 구현은 스토어 디렉터리에 직접 새 파일을 생성하므로 이름 변경 작업이 필요하지 않습니다. 각 저장소 디렉터리의 메타 파일을 기반으로 메모리에 커밋된 hfile 인스턴스 목록을 보관합니다. 새 hfile이 커밋될 때마다 지정된 저장소의 추적된 파일 목록이 업데이트되고 새 메타 파일이 목록 콘텐츠와 함께 작성되고 오래된 목록이 포함된 이전 메타 파일은 삭제됩니다. Store File Tracking에 대한 자세한 내용은 Apache HBase 참조 가이드의 Store File Tracking에서 확인할 수 있습니다.

FileBasedStoreFile 트래커 구현은 Amazon EMR 릴리스 7.4.0부터 기본적으로 활성화됩니다.

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.FileBasedStoreFileTracker"
}
```

FileBasedStoreFileTracker 구현을 비활성화하려면 클러스터를 시작할 때 다음 구성 재정의를 지정합니다.

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

Note

Amazon EMR 클러스터를 재구성할 때 모든 인스턴스 그룹을 업데이트해야 합니다.

운영상의 고려 사항

HBase 리전 서버는 BlockCache를 사용하여 데이터 읽기를 메모리에 저장하고 BucketCache를 사용하여 데이터 읽기를 로컬 디스크에 저장합니다. 또한 리전 서버는 MemStore를 사용하여 데이터 쓰기를 메모리에 저장하고, 데이터를 Amazon S3의 HBase StoreFiles에 쓰기 전에 write-ahead 로그를 사용하여 데이터 쓰기를 HDFS에 저장합니다. 클러스터의 읽기 성능은 인 메모리 또는 디스크 상의 캐시에서 레코드를 검색할 수 있는 빈도와 관련이 있습니다. 캐시가 누락되면 Amazon S3에 있는 StoreFile에서 레코드를 읽게 되고, 이 경우 HDFS에서 읽는 것보다 훨씬 더 높은 지연 시간과 표준 편차가 발생합니다. 또한 Amazon S3에 대한 최대 요청 속도가 로컬 캐시에서 얻을 수 있는 것보다 낮으므로 읽기가 많은 워크로드의 경우 데이터 캐싱이 중요할 수 있습니다. Amazon S3 성능에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 성능 최적화를 참조하세요.

성능을 향상시키려면 가능한 한 많은 데이터 세트를 EC2 인스턴스 스토리지에 캐시하는 것이 좋습니다. BucketCache는 리전 서버의 EC2 인스턴스 스토리지를 사용하기 때문에 충분한 인스턴스 스토어가 있는 EC2 인스턴스 유형을 선택하고 Amazon EBS 스토리지를 추가하여 필요한 캐시 크기를 수용

할 수 있습니다. `hbase.bucketcache.size` 속성을 사용하여 연결된 인스턴스 스토어 및 EBS 볼륨에서 BucketCache 크기를 늘릴 수도 있습니다. 기본 설정은 8,192MB입니다.

쓰기의 경우 MemStore 플러시의 빈도 및 사소한 압축 및 주요 압축 중에 제공되는 StoreFiles 수가 리전 서버 응답 시간 증가에 크게 기여할 수 있습니다. 최적의 성능을 얻기 위해 MemStore 플러시 및 HRegion 블록 승수 크기를 늘리면 주요 압축 사이의 경과 시간이 늘어나지만, 읽기 전용 복제본을 사용할 경우 일치하는 데 지연 시간이 길어질 수 있습니다. 경우에 따라 더 큰 파일 블록 크기(5GB 미만)를 사용하여 EMRFS에서 Amazon S3 멀티파트 업로드 기능을 트리거하여 성능을 향상시킬 수 있습니다. Amazon EMR의 기본 블록 크기는 128MB입니다. 자세한 내용은 HDFS 구성 단원을 참조하십시오. 플러시 및 압축으로 성능을 벤치마킹하는 동안 1GB 블록 크기를 초과하는 고객은 거의 없었습니다. 또한 HBase 압축 및 리전 서버는 더 적은 수의 StoreFile을 압축해야 할 때 최적으로 작동합니다.

큰 디렉터리는 이름을 바꾸어야 하기 때문에 테이블을 Amazon S3에서 삭제하는 데 상당한 시간이 걸릴 수 있습니다. 삭제하는 대신 테이블을 비활성화하는 것이 좋습니다.

오래된 WAL 파일과 스토어 파일을 정리하는 HBase 클리너 프로세스가 있습니다. Amazon EMR 릴리스 버전 5.17.0 이상을 사용하면 클리너가 전역적으로 활성화되고 다음 구성 속성을 사용하여 클리너 동작을 제어할 수 있습니다.

구성 속성	기본값	설명
`hbase.regionserver.hfilecleaner.large.thread.count`	1	만료된 큰 HFiles를 삭제하기 위해 할당된 스레드 수
`hbase.regionserver.hfilecleaner.small.thread.count`	1	만료된 작은 HFiles를 삭제하기 위해 할당된 스레드 수
`hbase.cleaner.scan.dir.concurrent.size`	사용 가능한 모든 코어의 1/4로 설정합니다.	oldWALs 디렉터리를 스캔하기 위한 스레드 수
`hbase.oldwals.cleaner.thread.size`	2	oldWALs 디렉터리 밑의 WALs를 삭제하기 위한 스레드 수

Amazon EMR 5.17.0 이전을 사용하는 경우 과중한 워크로드를 실행할 때 클리너 작업은 쿼리 성능에 영향을 줄 수 있으므로 사용량이 적은 시간에만 클리너를 사용하는 것이 좋습니다. 클리너에는 다음 HBase 셸 명령이 있습니다.

- `cleaner_chore_enabled`는 클리너가 활성화되었는지 여부를 쿼리합니다.
- `cleaner_chore_run`는 클리너를 수동으로 실행하여 파일을 제거합니다.
- `cleaner_chore_switch`는 클리너를 활성화 또는 비활성화하고 클리너의 이전 상태를 반환합니다. 예를 들어, `cleaner_chore_switch true`는 클리너를 활성화합니다.

Amazon S3 기반 HBase 성능 튜닝을 위한 속성

Amazon S3 기반 HBase를 사용할 때 다음 파라미터를 조정하여 워크로드 성능을 조정할 수 있습니다.

구성 속성	기본값	설명
`hbase.bucketcache.size`	8,192	리전 서버 Amazon EC2 인스턴스 스토어 및 BucketCache 스토리지용 EBS 볼륨에 예약된 디스크 공간(MB)입니다. 설정은 모든 리전 서버 인스턴스에 적용됩니다. 큰 BucketCache 크기는 일반적으로 향상된 성능에 해당합니다.
`hbase.hregion.memstore.flush.size`	134217728	Amazon S3에 대한 memstore 플러시가 트리거되는 데이터 한도(바이트)입니다.
`hbase.hregion.memstore.block.multiplier`	4	업데이트가 차단되는 MemStore 상한을 결정하는 승수입니다. MemStore가 이 값을 곱한 `hbase.hregion.memstore.flush.size` 를 초과하면 업데이트가 차단됩니다. MemStore 플러시 및 압축이 발생하여 업

구성 속성	기본값	설명
		데이트를 차단 해제할 수 있습니다.
`hbase.hstore.blockingStoreFiles`	10	업데이트가 차단되기 전에 스토어에 존재할 수 있는 StoreFiles의 최대 수입니다.
`hbase.hregion.max.filesize`	10737418240	리전을 분할하기 이전의 리전의 최대 크기입니다.

데이터 손실 없이 클러스터 종료 및 복원

Amazon S3에 기록되지 않은 데이터 손실 없이 Amazon EMR 클러스터를 종료하려면 MemStore 캐시를 Amazon S3에 플러시하여 새 스토어 파일을 작성해야 합니다. 먼저 모든 테이블을 비활성화해야 합니다. 다음 단계 구성은 클러스터에 단계를 추가할 때 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 AWS CLI 및 콘솔을 사용하여 단계 작업을 참조하세요.

```
Name="Disable all tables",Jar="command-runner.jar",Args=["/bin/bash","/usr/lib/hbase/bin/disable_all_tables.sh"]
```

또는 다음 bash 명령을 직접 실행할 수도 있습니다.

```
bash /usr/lib/hbase/bin/disable_all_tables.sh
```

모든 테이블을 비활성화한 후 HBase 셸과 다음 명령을 사용하여 `hbase:meta` 테이블을 플러시합니다.

```
flush 'hbase:meta'
```

그런 다음 Amazon EMR 클러스터에 제공된 쉘 스크립트를 실행하여 MemStore 캐시를 플러시할 수 있습니다. 단계별로 추가하거나 클러스터 내 AWS CLI를 사용하여 직접 실행할 수 있습니다. 이 스크립트는 모든 리전 서버의 MemStore가 Amazon S3로 플러시되도록 모든 HBase 테이블을 비활성화합니다. 스크립트가 성공적으로 완료되면 데이터는 Amazon S3에 유지되고 클러스터는 종료될 수 있습니다.

동일한 HBase 데이터로 클러스터를 다시 시작하려면에서 AWS Management Console 또는 `hbase.rootdir` 구성 속성을 사용하여 이전 클러스터와 동일한 Amazon S3 위치를 지정합니다.

Amazon EMR에 대한 미리 쓰기 로그(WAL)

Amazon EMR 6.15 이상을 사용하면 Amazon EMR WAL에 Apache HBase 미리 쓰기 로그(WAL)를 쓸 수 있습니다. Amazon EMR 릴리스가 낮으면 Amazon S3 기반 HBase 옵션을 사용하여 클러스터를 생성할 때 WAL은 클러스터용 로컬 디스크에 저장되는 유일한 Apache HBase 구성 요소이며, 루트 디렉터리, 저장소 파일(HFiles), 테이블 메타데이터 및 데이터와 같은 다른 구성 요소를 Amazon S3에 저장할 수 있습니다.

Amazon EMR WAL을 사용하여 Amazon S3로 플러시되지 않은 데이터를 복구할 수 있습니다. HBase 클러스터를 완전히 백업하려면 Amazon EMR WAL 서비스를 사용하도록 옵트인합니다. `RegionServer`는 백그라운드 작업으로 HBase 미리 쓰기 로그(WAL)를 Amazon EMR에 대한 WAL에 씁니다.

클러스터 또는 AZ가 비정상 상태이거나 사용할 수 없는 경우 새 클러스터를 생성하고 동일한 S3 루트 디렉터리와 Amazon EMR WAL 워크스페이스를 가리킨 후 몇 분 내에 WAL에서 데이터를 자동으로 복구할 수 있습니다. 자세한 내용은 [Amazon EMR WAL에서 복원](#) 단원을 참조하십시오.

Amazon EMR 릴리스 7.3.0 이상부터 Amazon EMR은 각 서버에 대해 여러 EMR WAL을 생성하고 여러 HBase 리전을 하나의 Amazon EMR WAL로 그룹화합니다. 이렇게 하면 Apache HBase WAL이 향상되어 로그 사용률을 개선하고 비용을 최적화합니다. HBase `RegionServer`당 Amazon EMR WAL 인스턴스 수를 구성하려면 `hbase.wal.regiongrouping.numgroups` 파라미터를 사용합니다. 이 파라미터는 기본적으로 2로 설정됩니다. WAL 그룹에는 메타 및 마스터 스토어라는 두 개의 시스템 테이블이 포함되지 않습니다. 이러한 테이블은 항상 자체 개별 WALs 사용합니다.

Amazon EMR 7.3.0보다 낮은 릴리스를 실행하는 경우 Amazon EMR WAL의 모든 데이터가 Amazon S3로 플러시되도록 이전 HBase 클러스터의 테이블을 수동으로 비활성화하는 것이 좋습니다. 그런 다음, 이전 Amazon EMR WAL을 삭제하고 이전 클러스터를 종료한 다음, 최신 릴리스를 실행하는 새 클러스터를 설정합니다. 문제가 발생하여 이전 클러스터에서 테이블을 비활성화할 수 없는 경우 이전 클러스터를 직접 종료하고 새 클러스터에서 `emr.wal.multiplex.migrate`를 `true`로 설정할 수 있습니다. true로 설정하면 HBase는 HBase 리전 초기화 중에 이전 Amazon EMR WAL 인스턴스의 데이터를 재생하고 재생 후 이전 WAL을 삭제하려고 시도합니다. 이 재생 프로세스에서는 읽기에 대한 추가 비용이 발생합니다. 마이그레이션 후에는 클러스터를 구성하고 `emr.wal.multiplex.migrate`를 `false`로 설정하는 것이 좋습니다. 또는 파라미터를 제거하여 HBase 리전 초기화 속도를 높일 수 있습니다.

> **Note**
>
> Amazon EMR WAL은 HBase가 데이터를 플러시한 후 데이터를 삭제합니다. HBase가 데이터를 플러시하지 않으면 Amazon EMR WAL은 최대 30일 동안 데이터를 유지합니다. 30일 후 Amazon EMR WAL은 데이터를 자동으로 삭제합니다. Amazon EMR은 EMR 클러스터를 종료한 후 최대 30일 동안 WAL 인스턴스를 보관합니다. 그러나 30일 이내에 동일한 S3 루트 디렉터리에서 새 WAL 지원 클러스터를 시작하면 Amazon EMR은 이전 클러스터에서 WAL 인스턴스를 삭제하지 않습니다. 자세한 내용은 Amazon EMR WAL에서 복원 단원을 참조하십시오.

다음 섹션에서는 HBase 지원 EMR 클러스터에서 Amazon EMR WAL을 설정하고 사용하는 방법을 설명합니다.

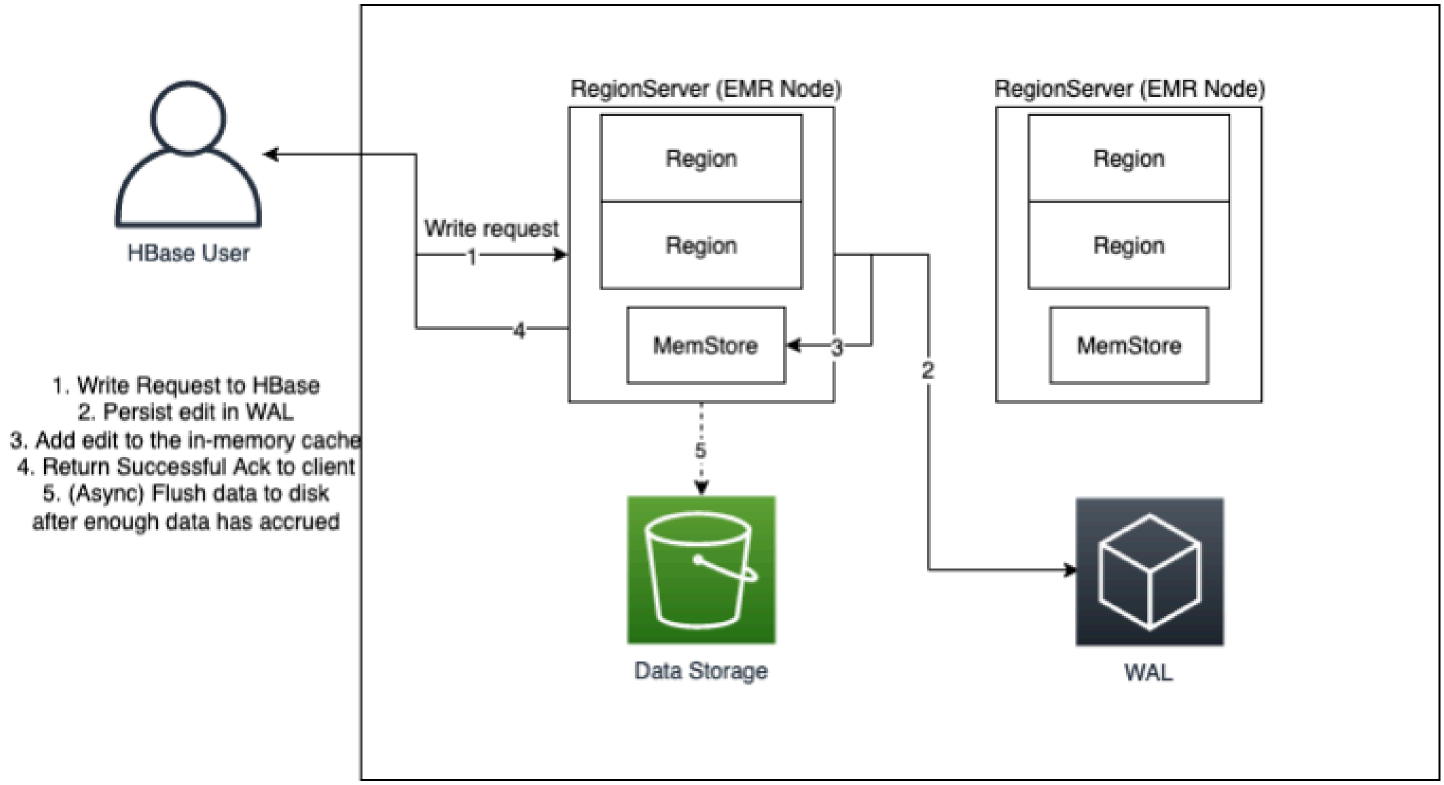
주제

- Amazon EMR WAL 워크스페이스
- Amazon EMR WAL에 필요한 권한
- Amazon EMR WAL 활성화
- Amazon EMR WAL에서 복원
- Amazon EMR WAL을 사용하여 보안 구성 사용
- AWS PrivateLink를 통해 Amazon EMR WAL에 액세스
- Amazon EMR WAL 요금 및 지표 이해
- WAL 워크스페이스 태그 지정
- EMR WAL 클러스터 간 복제
- Amazon EMR WAL에 대한 고려 사항 및 리전
- Amazon EMR WAL(EMRWAL) CLI 참조

Amazon EMR WAL 워크스페이스

Amazon EMR WAL은 WAL 워크스페이스의 개념을 추가합니다. WAL 워크스페이스는 WAL의 논리적 컨테이너입니다. Amazon EMR WAL의 각 미리 쓰기 로그는 WAL 워크스페이스에 의해 캡슐화됩니다. EMR 클러스터는 클러스터 시작 시 구성한 정확히 하나의 WAL 워크스페이스 또는 워크스페이스를 지정하지 않은 경우 `defaultWALworkspace`에 WAL을 씁니다. WAL 워크스페이스는 네임스페이스와 같은 기존 HBase 용어와 관련이 없습니다.

WAL 워크스페이스를 사용하여 클러스터가 액세스해야 하는 워크스페이스만 포함하도록 Amazon EMR WAL IAM 권한 범위를 축소할 수 있습니다. 태그 기반 액세스 제어를 위해 WAL 워크스페이스에 태그를 지정할 수도 있습니다. 태그 지정에 대한 자세한 내용은 WAL 워크스페이스 태그 지정 섹션을 참조하세요.



Amazon EMR WAL에 필요한 권한

클러스터를 Amazon EMR WAL에 연결하려면 클러스터의 인스턴스 프로파일에 특정 IAM 권한이 필요합니다.

- Amazon EMR WAL은 AWSServiceRoleForEMRWAL 서비스 연결 역할을 사용하여 클러스터 상태를 검색합니다. Amazon EMR은 WAL 워크스페이스를 생성할 때 이 서비스 연결 역할을 자동으로 생성합니다. 또는 Amazon EMR WAL에 대한 워크스페이스를 구성하고 서비스 연결 역할이 아직 존재하지 않을 때 HBase가 서비스 연결 역할을 생성합니다.

  클러스터에 대해 Amazon EMR WAL을 활성화하려면 먼저 AWSServiceRoleForEMRWAL 서비스 연결 역할의 자동 생성을 허용하도록 권한을 구성해야 합니다. 이 기능을 추가하는 예제 명령문과 자세한 내용은 미리 쓰기 로깅을 위해 서비스 연결 역할 사용을 참조하세요.
- Amazon EMR WAL은 HBase 미리 쓰기 로그(WAL)를 사용하기 때문에 클러스터는 HBase WAL을 사용해야 합니다. 다음은 HBase를 실행해야 하는 최소 IAM 권한입니다.다. 인스턴스 프로파일의 권한 정책에 다음을 추가합니다.

```
emrwal:DeleteWal
emrwal:CreateWal
emrwal:CreateWorkspace
emrwal:AppendEdit
emrwal:ReplayEdits
emrwal:GetCurrentWalTime
emrwal:CompleteWalFlush
emrwal:ListWALs
emrwal:DescribeWAL
emrwal:TrimWAL
emrwal:ArchiveWAL
emrwal:ArchiveWALCheckPoint
```

> **Note**
>
> Amazon EMR WAL에 대한 권한 범위를 최소 세트로만 지정하는 경우 일부 EMRWAL CLI 명령에는 실행에 필요한 권한이 없습니다.

Amazon EMR WAL 활성화

다음 단계를 통해 AWS Command Line Interface를 사용하여 클러스터를 생성하는 경우 Amazon EMR WAL에 쓰기를 활성화합니다.

> **Note**
>
> 이미 실행 중인 클러스터에 대해 Amazon EMR WAL을 활성화할 수 없으며, 동일한 S3 루트 디렉터리로 두 클러스터를 시작할 수 없습니다. 자세한 내용은 Amazon EMR WAL에 대한 고려 사항 및 리전 단원을 참조하십시오.

1. Amazon EMR WAL 지원 클러스터를 생성하려면 먼저 클러스터에서 사용할 인스턴스 프로파일에 필요한 권한을 추가해야 합니다. 자세한 내용은 Amazon EMR WAL에 필요한 권한 단원을 참조하십시오.
2. AWS CLI에서 클러스터를 생성합니다. `--configurations` 옵션을 사용하여 아래 예제와 같이 `hbase.emr.wal.enabled` 속성을 지정하는 JSON 구성 객체를 제공합니다.

- Amazon S3의 루트 디렉터리 위치 및 스토리지 모드를 지정합니다. 사용자가 지정하는 Amazon S3 위치는 EMR 클러스터와 동일한 리전에 있어야 하지만, 한 번에 하나의 활성 클러스터만 S3에서 동일한 HBase 루트 디렉터리를 사용할 수 있습니다.
- 인스턴스 그룹 구성을 사용하여 클러스터를 생성합니다. Amazon EMR WAL을 인스턴스 플릿 구성과 함께 사용할 수 없습니다. 인스턴스 그룹을 포함하는 클러스터 생성에 대한 자세한 내용은 Amazon EMR 관리 안내서의 인스턴스 플릿이나 균일한 인스턴스 그룹을 사용하여 클러스터 생성을 참조하세요.
- 클러스터를 생성하는 콘솔 단계와를 사용하는 자세한 `create-cluster` 예제는 HBase를 사용하여 클러스터 생성을 AWS CLI참조하세요.

3. 새 클러스터에 대해 WAL을 활성화하려면 `hbase.emr.wal.enabled` 속성을 `true`로 설정합니다. 다음 명령에는 예제 구성 객체가 있는 JSON 스니펫이 포함되어 있습니다.

```
aws emr create-cluster --name "hbasewal" --release-label emr-6.x.y \
--applications Name=HBase --use-default-roles --ec2-attributes KeyName=myKey \
--instance-type m6i.xlarge --instance-count 1 --configurations hbase.json
$cat hbase.json
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

새로 생성된 클러스터에서 HBase가 온라인 상태이면 HBase는 자동으로 Amazon EMR WAL에 WAL 데이터를 쓰고 복구 목적으로 Amazon EMR WAL을 사용합니다.

Example 1: Amazon EMR WAL을 사용하는 EMR 클러스터 생성

```
[
```

```
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Example 2: 사용자 지정 WAL 워크스페이스를 사용하여 EMR 클러스터 생성

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

Amazon EMR WAL에서 복원

원래 클러스터의 Amazon EMR WAL은 30일 동안 보존되므로 해당 30일 기간에 새로 생성된 클러스터에 대해 WAL을 복원하고 재사용할 수 있습니다. 동일한 S3 루트 디렉터리에서 새 클러스터를 시작하면 Amazon EMR은 이전 클러스터에서 WAL 인스턴스를 유지합니다. 이 새 클러스터를 종료하면 종료 시점부터 30일 클럭이 재시작됩니다.

다음 절차를 사용하여 새 클러스터로 기존 WAL을 복원합니다. 이 프로세스는 Amazon EMR WAL이 활성화된 상태에서 원래 클러스터를 생성했다고 가정합니다.

1. WAL 지원 클러스터를 생성한 후 30일 이내에 원래 클러스터 AWS 리전 와 동일한에 새 클러스터를 생성합니다. 새 클러스터는 원래 클러스터가 생성된 동일한 리전 내 동일한 AZ 또는 다른 AZ에 있을 수 있습니다.

   Amazon S3에서 스토리지 모드 및 루트 디렉터리 위치를 지정하도록 객체 속성을 구성합니다. 사용자가 지정하는 Amazon S3 위치는 EMR 클러스터와 동일한 리전에 있어야 하지만, 한 번에 하나의 활성 클러스터만 S3에서 동일한 HBase 루트 디렉터리를 사용할 수 있습니다.

   클러스터를 생성하는 콘솔 단계와를 사용하는 자세한 `create-cluster` 예제는 HBase를 사용하여 클러스터 생성을 AWS CLI참조하세요.
2. 새 클러스터에 대해 기존 Amazon EMR WAL을 사용하려면 `hbase.emr.wal.enabled` 속성을 `true`로 설정합니다. 다음 JSON 스니펫은 예제 구성 객체를 보여줍니다.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

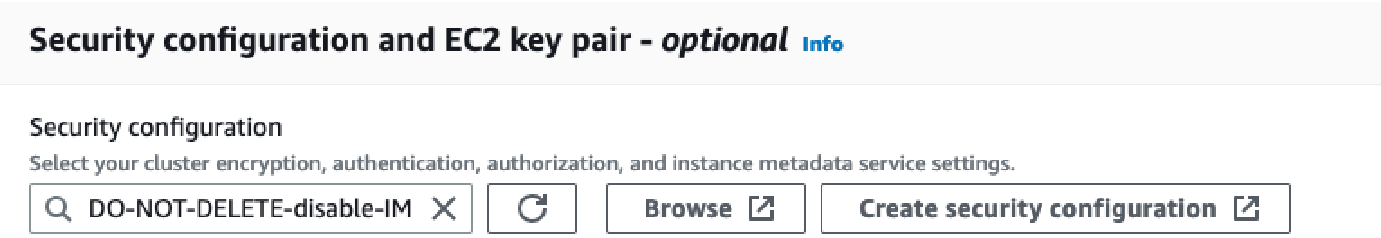
Amazon EMR WAL을 사용하여 보안 구성 사용

Amazon EMR은 클러스터와 Amazon EMR WAL 서비스 간에 전송 중 데이터와 Amazon EMR WAL의 저장 데이터를 모두 자동으로 암호화합니다. 자세한 내용은 Amazon EMR WAL에 대한 저장 데이터 암호화를 참조하세요. 보안 구성을 사용하여 AWS Key Management Service (KMS) 서비스에서 자체 키를 가져오고 Amazon EMR WAL에 저장하는 데이터를 암호화할 수도 있습니다.

클러스터를 생성하는 경우 다음 방법 중 하나를 사용하여 보안 구성을 선택합니다.

Console

의 보안 구성 및 EC2 키 페어에서 구성을 AWS Management Console지정합니다.



CLI

create-cluster 명령을 사용할 때에서 `--security-configuration` 파라미터를 AWS CLI설정합니다.

자세한 내용은 Amazon EMR 관리 안내서의 Amazon EMR WAL에 대한 저장 데이터 암호화 및 보안 구성을 사용하여 클러스터 보안 설정을 참조하세요.

WAL에 대한 자세한 보안 관련 정보는 미리 쓰기 로깅을 위해 서비스 연결 역할 사용을 참조하세요.

AWS PrivateLink를 통해 Amazon EMR WAL에 액세스

AWS 네트워크 내에서 연결을 유지하려면 Amazon EMR WAL에서 AWS PrivateLink 지원을 제공합니다. 설정하려면 AWS Management Console 또는 AWS Command Line Interface (AWS CLI)를 AWS PrivateLink사용하여 Amazon EMR WAL에 연결하는 인터페이스 VPC 엔드포인트를 생성합니다. 자세한 내용은 AWS PrivateLink 가이드의 인터페이스 VPC 엔드포인트를 사용하여 AWS 서비스 액세스를 참조하세요.

기본 단계는 다음과 같습니다.

1. Amazon VPC 콘솔을 사용하여 VPC 엔드포인트를 생성합니다. 엔드포인트를 선택한 다음, 엔드포인트 생성을 선택합니다.
2. 서비스 카테고리를 AWS 서비스로 유지합니다.
3. 서비스 패널의 검색창에서 **emrwal**을 입력하고 레이블이 `com.amazonaws.`*`region`*`.emrwal.prod`인 서비스를 선택합니다.
4. VPC를 선택하고 엔드포인트를 저장합니다. EMR 클러스터에 연결하는 VPC 엔드포인트에 동일한 보안 그룹을 연결해야 합니다.
5. 원하는 경우 이제 새 엔드포인트에 대한 프라이빗 DNS 호스트 이름을 활성화할 수 있습니다. VPC에 대해 DNS 호스트 이름 활성화 및 DNS 지원 활성화를 `true`로 설정합니다. 그런 다음, 엔

드포인트 ID를 선택하고 작업 메뉴에서 VPC 설정 편집을 선택한 다음, 프라이빗 DNS 이름을 활성화합니다.

- 엔드포인트의 프라이빗 DNS 호스트 이름은 `prod.emrwal.`*`region`*`.amazonaws.com` 형식을 따릅니다.
- 프라이빗 DNS 호스트 이름을 활성화하지 않는 경우 Amazon VPC는 *`endpointID`*`.prod.emrwal.`*`region`*`.vpce.amazonaws.com` 형식으로 DNS 엔드포인트 이름을 자동으로 제공합니다.

6. AWS PrivateLink 엔드포인트를 사용하려면 다음 예제와 같이 Amazon EMR WAL 지원 클러스터를 생성할 때 `emr.wal.client.endpoint` 구성을 수정합니다.

```
[
    {
        "Classification": "hbase-site",
        "Properties": {
            "hbase.rootdir": "s3://amzn-s3-demo-bucket/MyHBaseStore",
            "emr.wal.workspace": "customWorkspaceName",
            "emr.wal.client.endpoint": "https://prod.emrwal.region.amazonaws.com"
        }
    },
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.emr.storageMode": "s3",
            "hbase.emr.wal.enabled": "true"
        }
    }
]
```

VPCE 정책을 사용하여 Amazon EMR WAL API에 대한 액세스를 허용하거나 제한할 수도 있습니다. 자세한 정보는 AWS PrivateLink 가이드의 엔드포인트 정책을 사용하여 VPC 엔드포인트에 대한 액세스 제어를 참조하세요.

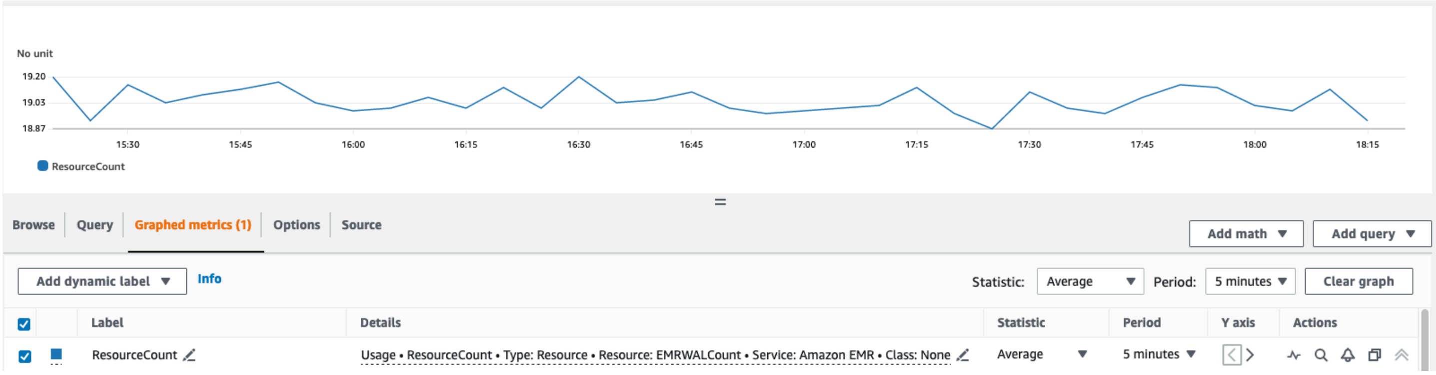
Amazon EMR WAL 요금 및 지표 이해

핵심 기능 청구 단위	세부 사항
EMR-WAL-Read-GiB	API calls to read data from your table are billed as ReadRequestGiB. This includes `Get and Scan`

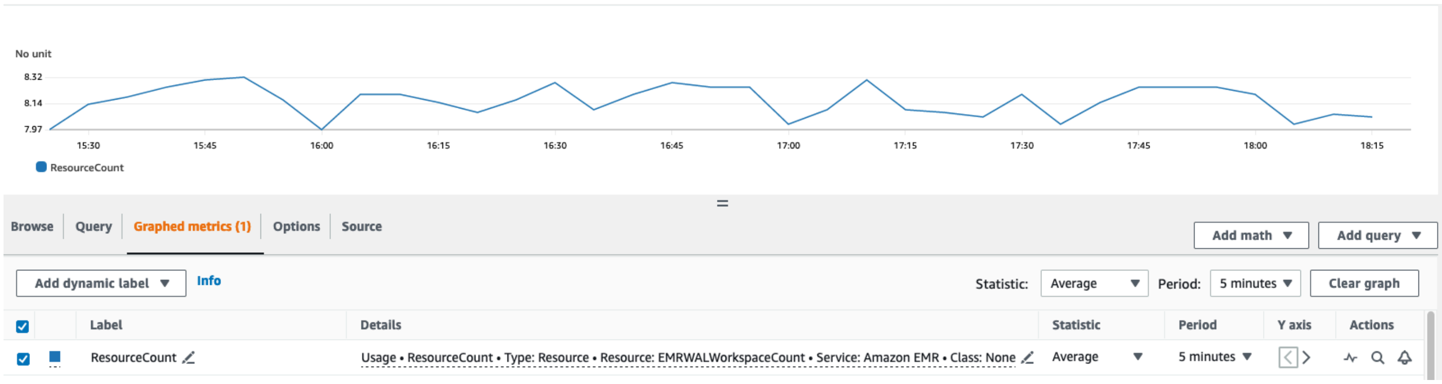
핵심 기능 청구 단위	세부 사항
	operations. Reads are charged based on the sizes of the read items. Amazon EMR bills at a minimum of 1 byte. For example, if you read a 1234.12 bytes item, you're charged for 1235 bytes. Reads are aggregated every hour for billing and shown as GiBs.
EMR-WAL-Write-GiB	API calls to write data from your table are billed as Write-GiB. This includes `Put` operations. Writes are charged based on the sizes of the written items. Amazon EMR bills at a minimum of 1 byte. For example, if you write a 1234.12 bytes item, you're charged for 1235 bytes. Writes are aggregated every hour for billing and shown as GiBs.

EMR-WAL-WALHours	서비스에 저장하는 WAL 수는 `EMR-WAL-WALHours` 로 청구됩니다. Amazon EMR에서는 HBase 리전당 하나의 WAL을 생성합니다. 예를 들어 시스템 테이블을 포함하여 HBase 테이블 20개를 생성하고 각 테이블에서 HBase 리전이 2개인 경우 다음과 같이 계산된 28,800시간의 WAL을 사용합니다.
	`  20 tables`
	`x  2 Regions per table`
	`x  1 WAL per Region`
	`x 30 days`
	`x 24 hours`
	`-----------`
	`28,800 EMR-WAL-WALHours`

예제 **`EMRWALCount`**:



예제 **`EMRWALWorkspaceCount`**:



WAL 워크스페이스 태그 지정

새 워크스페이스를 생성하는 경우 워크스페이스에 태그를 추가할 수 있으며 실행 중인 클러스터에 대한 활성 워크스페이스에서 태그를 추가, 제거 또는 나열할 수 있습니다. 워크스페이스의 개별 리소스에 태그를 지정할 수 없으며, 기존 태그를 업데이트할 수 없습니다. 대신 워크스페이스에서 원하지 않는 태그를 제거하고 바꿉니다.

EMRWAL CLI에서 워크스페이스에 태그를 지정할 수 있습니다. 워크스페이스에 태그를 지정하기 위한 EMRWAL CLI 명령 목록은 [Amazon EMR WAL(EMRWAL) CLI 참조](#) 섹션을 참조하세요.

다음 예제 IAM 정책에서는 적절한 태그 지정 키 `resource_tag_allow_test_key` 및 값 `resource_tag_allow_test_value`의 워크스페이스 CRUDL 작업만 허용하는 시나리오를 보여줍니다.

JSON

```
{
  "Version": "2012-10-17",
```

```
  "Statement": [
      {
          "Action": [
              "EMRWAL:*"
          ],
          "Effect": "Allow",
          "Resource": [
              "*",
              "*"
          ],
          "Condition": {
              "StringEquals": {
                  "aws:ResourceTag/resource_tag_allow_test_key": [
                      "resource_tag_allow_test_value"
                  ]
              }
          }
      }
  ]
}
```

이제 워크스페이스 작업에 태그가 필요한지 확인하려면 Amazon EMR WAL(EMRWAL) CLI 참조를 사용하여 원하는 리소스 태그가 있는 워크스페이스에 대한 `tagAllowResourceTag`에서 `listTagsForResource` 명령을 직접 호출합니다. 조건을 올바르게 구성하면 명령에 성공합니다.

```
emrwal listTagsForResource -r us-east-1 -arn arn:aws:emrwal:us-east-1:arn:workspace/
tagAllowResourceTag
Tag(Key=resource_tag_allow_test_key, Value=resource_tag_allow_test_value)
```

EMR WAL 클러스터 간 복제

EMR 7.5에서 EMR WAL은 미리 쓰기 로그의 HBase 클러스터 간 복제를 지원합니다. 이 주제에서는 기능을 활성화하고 작동 중인지 확인하는 방법을 보여줍니다. 클러스터 복제에 대한 자세한 내용은 Apache HBase 설명서의 클러스터 복제를 참조하세요.

> **Note**
>
> 복제 프로세스는 로컬 EMR WAL에서 데이터를 읽기 때문에 미리 쓰기 로그와 관련된 추가 읽기 비용이 발생합니다. 비용에 대한 자세한 내용은 Amazon EMR 릴리스 정보를 참조하세요.

클러스터 간 복제 설정

동일한 사용자 경험을 제공하기 위해 EMR WAL에서 복제 기능을 활성화하는 것은 기본 HBase 미리 쓰기 로그와 동일합니다. 다음 절차는 간단한 예제를 보여줍니다. 자세한 내용은 클러스터 복제의 HBase 설명서를 참조하세요.

1. EMR WAL이 활성화된 상태에서 복제 소스인 기본 클러스터를 시작합니다. 미리 쓰기 로그를 활성화하려면 섹션을 참조하세요Amazon EMR WAL 활성화. 또한 피어 클러스터를 시작합니다. 이 피어 클러스터의 경우 EMR WAL을 활성화할지 여부를 선택할 수 있습니다.
2. 두 클러스터 모두에서 테이블을 생성합니다.

```
HBASE_CMD="sudo -u hbase hbase"
echo "create 'test_replication_table',{NAME => 'CF'}" | $HBASE_CMD shell
```

3. 기본 클러스터에 피어 설정을 추가하고 테이블 복제를 활성화합니다. 피어를 추가하는 동안 피어 클러스터 마스터 노드 호스트 이름인 PEER_DNS가 필요합니다.

```
HBASE_CMD="sudo -u hbase hbase"
PEER_DNS="ip-10-1-1-0.ec2.com"
PEER_NAME="aws"
TABLE_NAME="test_replication_table"

## Create peering with the destination cluster
echo "add_peer '$PEER_NAME', CLUSTER_KEY => '$PEER_DNS:2181:/hbase'" | $HBASE_CMD
 shell

## List peers in the primary cluster to confirm peer setup
echo "list_peers" | $HBASE_CMD shell

## Enable table replication
echo "enable_table_replication '$TABLE_NAME'" | $HBASE_CMD shell
```

교차 클러스터 복제 확인

설정 단계를 수행하면 기본 클러스터와 피어 클러스터 간에 복제가 활성화됩니다. 복제가 작동하는지 확인하는 테스트가 이어집니다.

1. 기본 클러스터에 데이터를 추가하고 피어 클러스터에 복제된 데이터를 확인합니다.

```
## Write on primary cluster with HBase CLI
```

```
put 'test_replication_table', 'aaa', 'CF:a', 'aaa_a1'
put 'test_replication_table', 'bbb', 'CF:b', 'bbb_b1'
put 'test_replication_table', 'ccc', 'CF:c', 'ccc_c1'
```

2. 피어 클러스터에서 복제가 성공했는지 확인합니다. 이 경우 기본 클러스터에서 피어 클러스터로 기록된 복제된 데이터가 표시됩니다.

```
### Scan on peer cluster with HBase CLI

scan 'test_replication_table'
```

Amazon EMR WAL에 대한 고려 사항 및 리전

Amazon EMR WAL에 대한 고려 사항

다음 목록에서는 Amazon EMR WAL의 중요한 고려 사항 및 제한 사항을 설명합니다.

- Amazon EMR WAL은 Amazon EMR 릴리스 6.15.0 이상에서 사용할 수 있습니다.
- Amazon EMR WAL은 옵트인 유료 서비스입니다. 읽기, 쓰기 및 데이터 스토리지와 같은 사용 작업에 대한 비용을 지불합니다. 자세한 내용은 Amazon EMR WAL 요금 및 지표 이해 및 Amazon EMR 요금 페이지를 참조하세요.
- Amazon EMR WAL은 HBase 미리 쓰기 로그(WAL)를 사용합니다. Amazon EMR WAL을 사용하려면 클러스터에서 HBase WAL을 사용해야 합니다.
- 클러스터를 생성할 때 Amazon EMR WAL을 활성화하려면 필수 역할 권한이 있어야 합니다. 자세한 내용은 미리 쓰기 로깅에 대한 서비스 연결 역할 사용을 참조하세요.
- AWS Management Console, AWS CLI또는 API를 사용하여 클러스터를 생성할 때 Amazon EMR WAL을 활성화해야 하며 인스턴스 그룹 구성을 사용해야 합니다. Amazon EMR WAL을 사용하여 클러스터를 생성하지 않은 경우 실행 중인 클러스터에서 Amazon EMR WAL을 활성화할 수 없습니다. 또한 실행 중인 클러스터에서 Amazon EMR WAL을 활성화하도록 `hbase-site` 구성을 편집할 수 없습니다.
- 루트 디렉터리에서 Amazon S3를 사용하는 클러스터에서만 Amazon EMR WAL을 활성화할 수 있습니다.
- Amazon EMR 버전 7.5.0 이전에는 Amazon EMR WAL의 레코드가 4MB 이하여야 했습니다. 그러나 Amazon EMR 버전 7.5.0 이상에서는 속성을 사용하여 EMR WAL의 최대 레코드 크기를 구성할 수 있습니다`emr.wal.max.payload.size`. 기본값은 1GB입니다. 다음 예제에서는 최대 레코드 크기를 2GB로 설정합니다.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "emr.wal.max.payload.size": "2147483648"
    }
  }
]
```

- Amazon S3의 동일한 HBase 루트 디렉터리에 활성 클러스터가 여러 개 있을 수 없습니다.
- 읽기 전용 복제본 클러스터에서 Amazon EMR WAL을 활성화할 수 없습니다.
- WAL은 관리형 서비스 내 여러 가용 영역에서 복제됩니다.
- WAL은 클러스터의 수명을 초과하고 다음 클러스터에 계속 사용할 수 있습니다.
- 시작 중에 또는 클러스터가 작동 중일 때(실행 중 상태) Amazon EMR WAL을 비활성화할 수 없습니다.
- WAL 및 워크스페이스 제한에 대한 자세한 내용은 Amazon EMR 엔드포인트 및 할당량 섹션을 참조하세요.

Amazon EMR WAL에 대한 리전 가용성

Amazon EMR WAL 서비스는 AWS 리전다음에서 사용할 수 있습니다.

- `ap-northeast-1` – 아시아 태평양(도쿄)
- `ap-northeast-2` – 아시아 태평양(서울)
- `ap-southeast-1` – 아시아 태평양(싱가포르)
- `ap-south-1` – 아시아 태평양(뭄바이)
- `ap-southeast-2` – 아시아 태평양(시드니)
- `eu-central-1` – 유럽(프랑크푸르트)
- `eu-north-1` – 유럽(스톡홀름)
- `eu-west-1` – 유럽(아일랜드)
- `sa-east-1` – 남아메리카(상파울루)
- `us-east-1` - 미국 동부(버지니아 북부)
- `us-east-2` - 미국 동부(오하이오)
- `us-west-2` - 미국 서부(오레곤)

다음 리전은 Amazon EMR 버전 7.3.0 이상에서만 사용할 수 있습니다.

- `ap-east-1` – 아시아 태평양(홍콩)
- `af-south-1` – 아프리카(케이프타운)
- `ca-central-1` - 캐나다(중부)
- `eu-west-2` – 유럽(런던)

Amazon EMR WAL(EMRWAL) CLI 참조

EMRWAL 명령줄 인터페이스(EMRWAL CLI)는 Amazon EMR에 대한 미리 쓰기 로그(WAL)를 관리하는 통합 도구입니다. 클러스터를 생성할 때 WAL을 활성화하면 EMRWAL CLI는 EMR 클러스터와 함께 제공됩니다. WAL 활성화에 대한 자세한 내용은 Amazon EMR에 대한 미리 쓰기 로그(WAL) 섹션을 참조하세요.

EMRWAL CLI에는 다음 명령이 포함됩니다.

주제

`createWorkspace`

`createWorkspace` 명령은 새 Amazon EMR WAL 워크스페이스를 생성합니다.

사용량:

```
emrwal createWorkspace [-tags <tags>] [-e {endpoint}] [-r {Region}] -w {workspacename}
 [-h]
```

예:

```
emrwal createWorkspace -w examplews
```

deleteWal

deleteWals 명령은 사용자가 지정한 Amazon EMR WAL을 삭제합니다.

사용량:

```
emrwal deleteWal [-e {endpoint}] [-r {Region}] [-w {workspacename}] [-p <tablePrefix>]
 [-n <walName>] [-N <fullName>] [-R] [-m] [-h]
```

예:

```
emrwal deleteWal -w examplews -p hbasetable -n examplewal
```

deleteWorkspace

deleteWorkspace 명령은 사용자가 지정한 Amazon EMR WAL 워크스페이스를 삭제합니다.

사용량:

```
emrwal deleteWorkspace [-e {endpoint}] [-r {Region}] -w {workspacename} [-h]
```

예:

```
emrwal deleteWorkspace -w examplews
```

listTagsForResource

listTagsForResource 명령은 사용자가 지정한 Amazon EMR WAL 워크스페이스의 모든 키-값 페어 태그를 나열합니다.

사용량:

```
emrwal listTagsForResource -arn {resource-arn} [-e {endpoint}] [-r {Region}] [-h]
```

예:

```
emrwal listTagsForResource -arn arn:aws:emrwal::1234567891234:workspace/examplews
```

listWals

listWals 명령은 사용자가 지정한 워크스페이스의 모든 Amazon EMR WAL을 나열합니다.

사용량:

```
emrwal listWals [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}] [-
r {Region}] [-w {workspacename}] [-p <tablePrefix>] [-M {integer}] [-h]
```

예:

```
emrwal listWals -w examplews
```

listWorkspaces

이 listWorkspaces 명령은 사용 가능한 모든 Amazon EMR WAL 워크스페이스를 나열합니다.

사용량:

```
emrwal listWorkspaces [-nextToken {token-string}] [-pageSize {integer}] [-e {endpoint}]
 [-r {Region}] [-M {integer}] [-h]
```

예:

```
emrwal listWorkspaces
```

tagResource

tagResource 명령은 사용자가 지정한 Amazon EMR WAL 워크스페이스에 하나 이상의 키-값 페어 태그를 할당합니다.

사용량:

```
emrwal tagResource -arn {resource-arn} -tags <tags> [-e {endpoint}] [-r {Region}] [-h]
```

예:

```
emrwal tagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tags tag_key=tag_value
```

untagResource

`untagResource` 명령은 사용자가 지정한 Amazon EMR WAL 워크스페이스에 하나 이상의 키-값 페어 태그를 할당 해제합니다.

사용량:

```
emrwal untagResource -arn {resource-arn} -tagKeys <tagKeys> [-e {endpoint}] [-
r {Region}] [-h]
```

예:

```
emrwal untagResource -arn arn:aws:emrwal::1234567891234:workspace/examplews -
tagKeys tag_key
```

HBase 쉘 사용

HBase 클러스터를 생성한 후 다음 단계는 데이터(읽기 전용 복제본 클러스터에서 지원되지 않는 데이터) 읽기 및 쓰기를 시작할 수 있도록 HBase에 연결하는 것입니다. HBase 셸을 사용하여 명령을 테스트할 수 있습니다.

HBase 셸을 열려면

1. SSH를 사용하여 HBase 클러스터의 기본 서버에 연결합니다. SSH를 사용하여 프라이머리 노드에 연결하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서의 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. `hbase shell`을(를) 실행합니다. 다음 예제와 유사한 프롬프트와 함께 HBase 쉘이 열립니다.

   ```
   hbase(main):001:0>
   ```

프롬프트에서 HBase 셸 명령을 실행할 수 있습니다. 셸 명령 및 명령 호출 방법에 대한 자세한 내용을 보려면 HBase 프롬프트에서 help를 입력하고 Enter 키를 누르십시오.

테이블 생성

다음 명령은 'f1'이라는 단일 열 패밀리가 있는 't1'이라는 테이블을 생성합니다.

```
hbase(main):001:0>create 't1', 'f1'
```

값 입력

다음 명령은 테이블 't1' 및 열 'f1'의 행 'r1'에 대해 값 'v1'을 입력합니다.

```
hbase(main):001:0>put 't1', 'r1', 'f1:col1', 'v1'
```

값 가져오기

다음 명령은 테이블 't1'의 행 'r1'에 대한 값을 가져옵니다.

```
hbase(main):001:0>get 't1', 'r1'
```

테이블 삭제

다음 명령은 테이블 't1'을 삭제합니다.

```
hbase(main):001:0>drop 'ns1:t1',false
```

부울 값은 테이블의 아카이브 여부에 해당하므로 저장하려는 경우 이 값을 `true`로 설정할 수 있습니다. 또한 부울 없이 `drop 'ns1:t1'`을 실행하여 테이블을 아카이브할 수도 있습니다.

Hive를 사용하여 HBase 테이블 액세스

HBase 및 [Apache Hive](https://hive.apache.org/)는 긴밀하게 통합되어 있으므로 HBase에 저장된 데이터에서 대량 병렬 처리 워크로드를 직접 실행할 수 있습니다. Hive를 HBase와 함께 사용하려면 대개 같은 클러스터에서 이 둘을 실행합니다. 그러나 별도의 클러스터에서 Hive 및 HBase를 실행할 수 있습니다. 다른 클러스터에서 HBase 및 Hive를 실행하면 애플리케이션이 클러스터 리소스를 더 효율적으로 사용할 수 있기 때문에 성능이 개선될 수 있습니다.

다음 절차에서는 Hive를 사용하여 클러스터의 HBase에 연결하는 방법을 보여줍니다.

> **Note**
>
> Hive 클러스터는 단일 HBase 클러스터에만 연결할 수 있습니다.

Hive를 HBase에 연결하려면

1. Hive 및 HBase가 설치된 별도의 클러스터를 생성하거나, HBase 및 Hive가 모두 설치된 단일 클러스터를 생성합니다.
2. 별도의 클러스터를 사용하는 경우 HBase 및 Hive 포트가 이 두 프라이머리 노드 사이에서 열리도록 보안 그룹을 수정합니다.
3. SSH를 사용하여 Hive가 설치된 클러스터의 프라이머리 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [SSH를 사용하여 프라이머리 노드에 연결]을 참조하세요.
4. 다음 명령을 사용하여 Hive 셸을 시작하십시오.

```
hive
```

5. (선택 사항) HBase 및 Hive가 동일한 클러스터에 있는 경우 이를 수행할 필요가 없습니다. Hive 클러스터의 HBase 클라이언트를 데이터를 포함하고 있는 HBase 클러스터에 연결합니다. 다음 예제에서는 *`public-DNS-name`*이 HBase 클러스터의 프라이머리 노드에 대한 퍼블릭 DNS 이름으로 바뀝니다(예: `ec2-50-19-76-67.compute-1.amazonaws.com`).

```
set hbase.zookeeper.quorum=public-DNS-name;
```

6. 필요에 따라 HBase 데이터에서 Hive 쿼리를 계속 실행하거나 다음 절차를 참조하십시오.

Hive에서 HBase 데이터에 액세스하려면

- 이전 절차에서처럼 Hive 및 HBase 클러스터 간의 연결이 설정된 후 Hive에 외부 테이블을 만들어서 HBase 클러스터에 저장된 데이터에 액세스할 수 있습니다.

  프라이머리 노드의 Hive 프롬프트에서 실행하는 경우 다음 예제는 `inputTable`이라는 HBase 테이블에 저장된 데이터를 참조하는 외부 테이블을 생성합니다. 그러면 Hive 명령문의 `inputTable`을 참조하여 HBase 클러스터에 저장된 데이터를 쿼리 및 수정할 수 있습니다.

```
set hbase.zookeeper.quorum=ec2-107-21-163-157.compute-1.amazonaws.com;

create external table inputTable (key string, value string)
     stored by 'org.apache.hadoop.hive.hbase.HBaseStorageHandler'
      with serdeproperties ("hbase.columns.mapping" = ":key,f1:col1")
      tblproperties ("hbase.table.name" = "t1");

select count(key) from inputTable ;
```

HBase와 Hive를 결합한 고급 사용 사례와 예제는 AWS 빅 데이터 블로그 게시물인 Amazon EMR에서 Apache HBase와 Apache Hive를 사용하여 NoSQL과 대규모 병렬 분석을 결합하기를 참조하세요.

HBase 스냅샷 사용

HBase는 기본 제공된 스냅샷 기능을 사용하여 테이블의 간소화된 백업을 생성합니다. EMR 클러스터에서 EMRFS를 사용하여 이 백업을 Amazon S3로 내보낼 수 있습니다. HBase 셸을 사용하여 프라이머리 노드에서 스냅샷을 생성할 수 있습니다. 이 주제에서는 쉘과 대화식으로 또는 `command-runner.jar` AWS CLI 또는와 함께를 사용하여 단계를 통해 이러한 명령을 실행하는 방법을 보여줍니다 AWS SDK for Java. 다른 유형의 HBase 백업에 대한 자세한 내용은 HBase 설명서에서 HBase backup을 참조하세요.

테이블을 사용하여 스냅샷 생성

```
hbase snapshot create -n snapshotName -t tableName
```

AWS CLI다음에서 command-runner.jar 사용:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "create","-n","snapshotName","-t","tableName"]
```

AWS SDK for Java

```
HadoopJarStepConfig hbaseSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","create","-n","snapshotName","-t","tableName");
```

> **Note**
>
> 스냅샷 이름이 고유하지 않으면 생성 작업이 -1 또는 255 반환 코드와 함께 실패합니다. 그러나 오류 내용을 설명하는 오류 메시지는 표시되지 않을 수 있습니다. 동일한 스냅샷 이름을 사용하려면 스냅샷을 삭제한 다음, 다시 생성합니다.

스냅샷 삭제

```
hbase shell
>> delete_snapshot 'snapshotName'
```

스냅샷 정보 보기

```
hbase snapshot info -snapshot snapshotName
```

Amazon S3로 스냅샷 내보내기

> **Important**
>
> 스냅샷을 내보낼 때 `-mappers` 값을 지정하지 않은 경우 HBase에서는 임의의 계산을 사용하여 매퍼 수를 결정합니다. 테이블의 크기에 따라 이 값이 매우 클 수 있습니다. 그러면 내보내기 중에 실행 중인 작업에 악영향을 주게 됩니다. 따라서 `-mappers` 파라미터, `-bandwidth` 파라미터(초당 메가바이트 단위로 대역폭 사용 지정) 또는 둘 모두를 지정하여 내보내기 작업에서 사용되는 클러스터 리소스를 제한하는 것이 좋습니다. 또는 사용량이 적은 기간 동안 스냅샷 내보내기 작업을 실행할 수 있습니다.

```
hbase snapshot export -snapshot snapshotName \
-copy-to s3://amzn-s3-demo-bucket/folder -mappers 2
```

`command-runner.jar`에서 사용 AWS CLI:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar",\
Args=[ "hbase", "snapshot", "export","-snapshot","snapshotName","-copy-to","s3://amzn-s3-demo-bucket/folder","-mappers","2","-bandwidth","50"]
```

AWS SDK for Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("hbase","snapshot","export",
      "-snapshot","snapshotName","-copy-to",
      "s3://bucketName/folder",
      "-mappers","2","-bandwidth","50");
```

Amazon S3에서 스냅샷 가져오기

가져오기 작업이지만 여기서 사용된 HBase 옵션은 여전히 `export`입니다.

```
sudo -u hbase hbase snapshot export \
-D hbase.rootdir=s3://amzn-s3-demo-bucket/folder \
-snapshot snapshotName \
-copy-to hdfs://masterPublicDNSName:8020/user/hbase \
-mappers 2
```

`command-runner.jar`에서 사용 AWS CLI:

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Name="HBase Shell Step",Jar="command-runner.jar", \
Args=["sudo","-u","hbase","hbase snapshot export","-snapshot","snapshotName", \
"-D","hbase.rootdir=s3://amzn-s3-demo-bucket/folder", \
"-copy-to","hdfs://masterPublicDNSName:8020/user/hbase","-mappers","2","-chmod","700"]
```

AWS SDK for Java:

```
HadoopJarStepConfig hbaseImportSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
  .withArgs("sudo","-u","hbase","hbase","snapshot","export", "-D","hbase.rootdir=s3://
path/to/snapshot",
      "-snapshot","snapshotName","-copy-to",
      "hdfs://masterPublicDNSName:8020/user/hbase",
      "-mappers","2","-chuser","hbase");
```

HBase 셸 내의 스냅샷에서 테이블 복원

```
hbase shell
```

```
>> disable tableName
>> restore_snapshot snapshotName
>> enable tableName
```

HBase는 현재 HBase 셸에 있는 모든 스냅샷 명령을 지원하지 않습니다. 예를 들어, 스냅샷을 복원하는 HBase 명령줄 옵션이 없기 때문에 셸 내에서 스냅샷을 복원해야 합니다. 즉 `command-runner.jar`에서 Bash 명령을 실행해야 합니다.

> **Note**
>
> 여기서 사용된 명령이 echo이기 때문에 Amazon EMR에서 실행한 명령이 `0` 종료 코드를 반환할 경우에도 쉘 명령이 계속 실패할 수 있습니다. 셸 명령을 단계로 실행하도록 선택한 경우 단계 로그를 확인하십시오.

```
echo 'disable tableName; \
restore_snapshot snapshotName; \
enable tableName' | hbase shell
```

다음은 AWS CLI를 사용한 단계입니다. 먼저 다음 `snapshot.json` 파일을 생성합니다.

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo $'disable \"tableName\"; restore_snapshot
 \"snapshotName\"; enable \"tableName\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./snapshot.json
```

AWS SDK for Java:

```
HadoopJarStepConfig hbaseRestoreSnapshotConf = new HadoopJarStepConfig()
  .withJar("command-runner.jar")
```

```
  .withArgs("bash","-c","echo $'disable \"tableName\"; restore_snapshot \"snapshotName
\"; enable \"snapshotName\"' | hbase shell");
```

HBase 구성

대부분의 애플리케이션에서는 기본 HBase 설정으로도 충분하지만, HBase 구성 설정을 수정할 수 있습니다. 이렇게 하려면 HBase 구성 분류의 속성을 사용하십시오. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

다음 예제에서는 Amazon S3에 저장된 `myConfig.json` 구성 파일에 따라 대체 HBase 루트 디렉터리가 있는 클러스터를 생성합니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=HBase \
--instance-type m5.xlarge --instance-count 3 --configurations https://s3.amazonaws.com/
amzn-s3-demo-bucket/myfolder/myConfig.json
```

다음 예제에 표시된 대로 `myConfig.json` 파일은 `hbase.rootdir` 구성 분류에 대한 `hbase-site` 속성을 지정합니다. *ip-XXX-XX-XX-XXX.ec2.internal*을 클러스터 프라이머리 노드의 내부 DNS 호스트 이름으로 바꿉니다.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.rootdir": "hdfs://ip-XXX-XX-XX-XXX.ec2.internal:8020/user/
myCustomHBaseDir"
    }
  }
]
```

Note

Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

YARN의 메모리 할당 변경

HBase가 YARN 애플리케이션으로 실행되고 있지 않으므로 YARN 및 해당 애플리케이션에 할당된 메모리를 다시 계산해야 합니다. 그러면 HBase가 설치된 경우 YARN에서 사용할 수 있는 전체 메모리가 줄어듭니다. 동일한 클러스터에서 YARN 애플리케이션과 HBase를 함께 배치할 때 이를 고려해야 합니다. 64GB보다 적은 메모리를 가진 인스턴스 유형은 NodeManager에서 사용할 수 있는 메모리의 절반을 차지하고, HBase RegionServer에 할당됩니다. 예를 들어, 메모리가 64GB보다 큰 인스턴스 유형의 경우 HBase RegionServer 메모리는 32GB로 제한됩니다. 일반적으로 YARN 설정 메모리는 MapReduce reducer 작업 메모리의 여러 배입니다.

작업 구성 설정의 기본값의 테이블은 HBase에 필요한 메모리를 기반으로 YARN 설정이 변경되었음을 보여줍니다.

HBase 포트 번호

HBase용으로 선택된 일부 포트 번호는 기본값과 다릅니다. Amazon EMR의 HBase에 대한 인터페이스와 포트는 다음과 같습니다.

HBase 포트

인터페이스	Port	프로토콜
HMaster	16000	TCP
HMaster UI	16010	HTTP
RegionServer	16020	TCP
RegionServer 정보	16030	HTTP
REST 서버	8070	HTTP

인터페이스	Port	프로토콜
REST UI	8085	HTTP
Thrift 서버	9090	TCP
Thrift 서버 UI	9095	HTTP

Important

Amazon EMR 릴리스 버전 4.6.0 이상에서 `kms-http-port`는 9700이며 `kms-admin-port`는 9701입니다.

최적화할 HBase 사이트 설정

애플리케이션의 워크로드에 맞게 HBase 클러스터를 최적화하도록 HBase 사이트 설정 중 일부 또는 전부를 설정할 수 있습니다. 다음 설정을 조사의 시작점으로 삼는 것이 좋습니다.

zookeeper.session.timeout

기본 제한 시간은 40초(40000밀리초)입니다. 리전 서버가 충돌할 경우, 이는 마스터 서버가 리전 서버의 부재를 감지하고 복구를 시작하는 데 걸리는 시간입니다. 마스터 서버의 복구 속도를 높이려면 이 값을 더 짧은 기간으로 줄이십시오. 다음 예에서는 30초 또는 30000밀리초를 사용합니다.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "zookeeper.session.timeout": "30000"
    }
  }
]
```

hbase.regionserver.handler.count

리전 서버가 테이블에 대한 요청을 처리하기 위해 열고 있는 스레드의 수를 정의합니다. 동시 클라이언트의 수가 많은 대형 쓰기 버퍼를 사용할 때 사용자가 리전 서버를 종료하지 못하게 하려면 기본값 10이 낮습니다. 요청당 페이로드가 MB 범위에 가까울 때(빅 puts, 라지 캐시를 사용한 스캔)는 이 값을 낮

게 유지하고, 페이로드가 작을 때(gets, 스몰 puts, ICVs, deletes)는 이 값을 높게 유지하는 것이 경험적 규칙입니다. 다음 예제에서는 열린 스레드의 수를 30으로 늘립니다.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.regionserver.handler.count": "30"
    }
  }
]
```

hbase.hregion.max.filesize

이 파라미터는 개별 리전의 크기(바이트)를 제어합니다. 기본적으로 1073741824로 설정됩니다. HBase 클러스터에 많은 양의 데이터를 쓰고 있어서 분할이 자주 발생하는 경우 이 크기를 늘려 개별 리전을 더 크게 만들 수 있습니다. 분할은 줄이지만 한 서버에서 다른 서버로 리전 로드를 분산하는 데 더 많은 시간이 필요합니다.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.hregion.max.filesize": "1073741824"
    }
  }
]
```

hbase.hregion.memstore.flush.size

이 파라미터는 디스크로 플러시되기 전에 memstore의 최대 크기를 바이트 단위로 관리합니다. 기본 설정은 134217728입니다. 워크로드가 짧은 쓰기 작업으로 구성되는 경우 이 제한을 늘려 모든 쓰기가 버스트 기간 동안 메모리에 남아있게 하고 나중에 디스크로 플러시되도록 할 수 있습니다. 이 경우 버스트 동안 성능을 향상시킬 수 있습니다.

```
[
  {
    "Classification":"hbase-site",
    "Properties": {
       "hbase.hregion.memstore.flush.size": "134217728"
```

```
    }
  }
]
```

성능 고려 사항

HBase용 ZGC(Z Garbage Collector) 활성화

Amazon EMR 버전 7.10.0 이상을 사용하면 사용자는 HBase 클러스터에 대한 가비지 수집(GC) 설정을 원활하게 구성할 수 있습니다. 지연 시간이 짧고 GC 일시 중지 시간이 밀리초 미만인 경우 ZGC(Z Garbage Collector)를 활성화하는 것이 좋습니다.

다음은 HBase RegionServer(들)에 대해 ZGC를 활성화하는 구성입니다.

```
[
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
                    "JAVA_HOME": "/usr/lib/jvm/jre-21",
                    "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational\""
                }
            }
        ]
    }
]
```

> **Note**
>
> JAVA_HOME 환경 변수는 세대 ZGC(권장)를 사용할 때 JDK 21에 도입되었으므로 설정해야 합니다. ZGC의 비세대 모드(제외`-XX:+ZGenerational`)를 사용하려면 JAVA_HOME을 설정할 필요가 없습니다. JDK 24에서는 ZGC의 비세대 모드가 제거되었습니다.

ZGC 튜닝

1. JVM 고정 힙 활성화

ZGC(Z Garbage Collector)는 고정 힙 메모리로 구성할 때 더 효율적으로 작동하므로 운영 체제에 메모리를 반환하는 오버헤드가 발생하지 않습니다. HBase 클러스터에 고정 힙 메모리를 구성하려면 다음 구성을 사용합니다.

```
[
    {
        "Classification": "hbase",
        "Properties": {
            "hbase.regionserver.fixed.heap.enabled": "true"
        }
    }
]
```

2. 사전 터치 활성화

프리터치를 활성화하면 초기 지연 시간을 줄이고 보다 예측 가능한 성능을 제공하여 가비지 수집(GC) 성능이 향상됩니다. HBase 클러스터에 대한 사전 터치를 활성화하려면 다음 구성을 사용합니다.

```
[
    {
        "Classification": "hbase-env",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "export",
                "Properties": {
                    "JAVA_HOME": "/usr/lib/jvm/jre-21",
                    "HBASE_REGIONSERVER_GC_OPTS": "\"-XX:+UseZGC -XX:+ZGenerational -
XX:+AlwaysPreTouch\""
                }
            }
        ]
    }
]
```

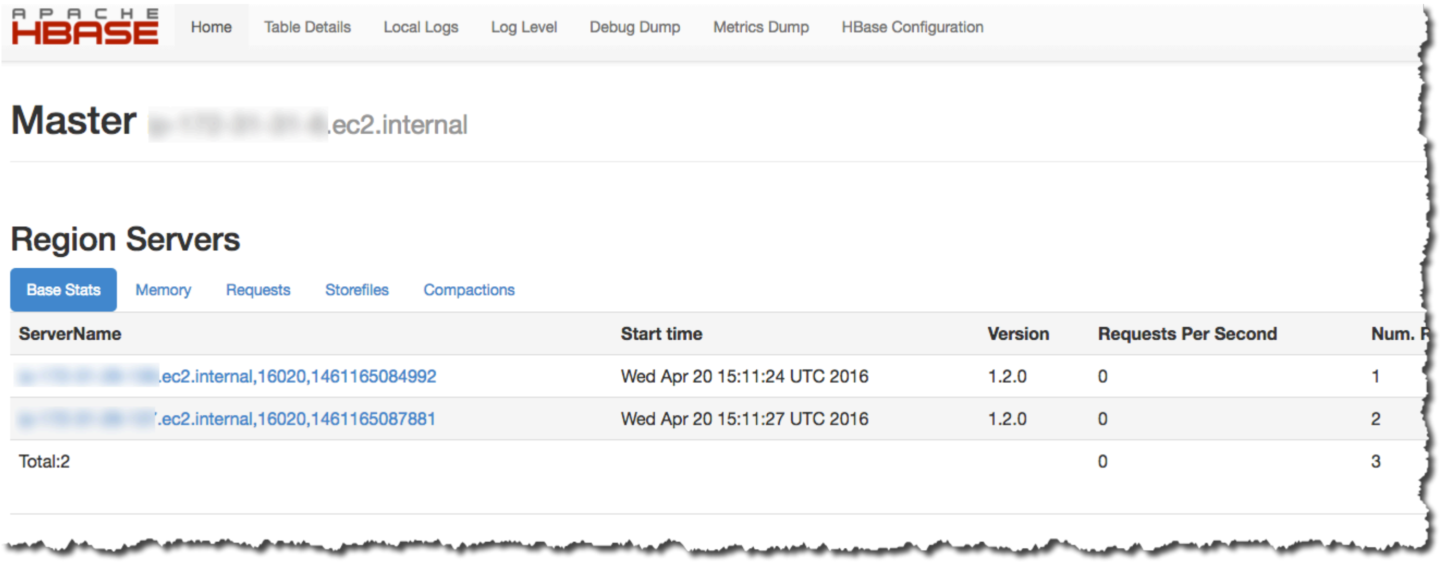
HBase 사용자 인터페이스 보기

> **Note**
>
> HBase 사용자 인터페이스는 기본적으로 비보안 HTTP 연결을 사용합니다. 보안 HTTP(HTTPS)를 활성화하려면 HBase 구성에서 `hbase-site` 분류에 대한 `hbase.ssl.enabled` 속성을 `true`로 설정합니다. HBase 웹 UI에서 보안 HTTP(HTTPS)를 사용하는 방법에 대한 자세한 내용은 Apache HBase Reference Guide를 참조하세요.

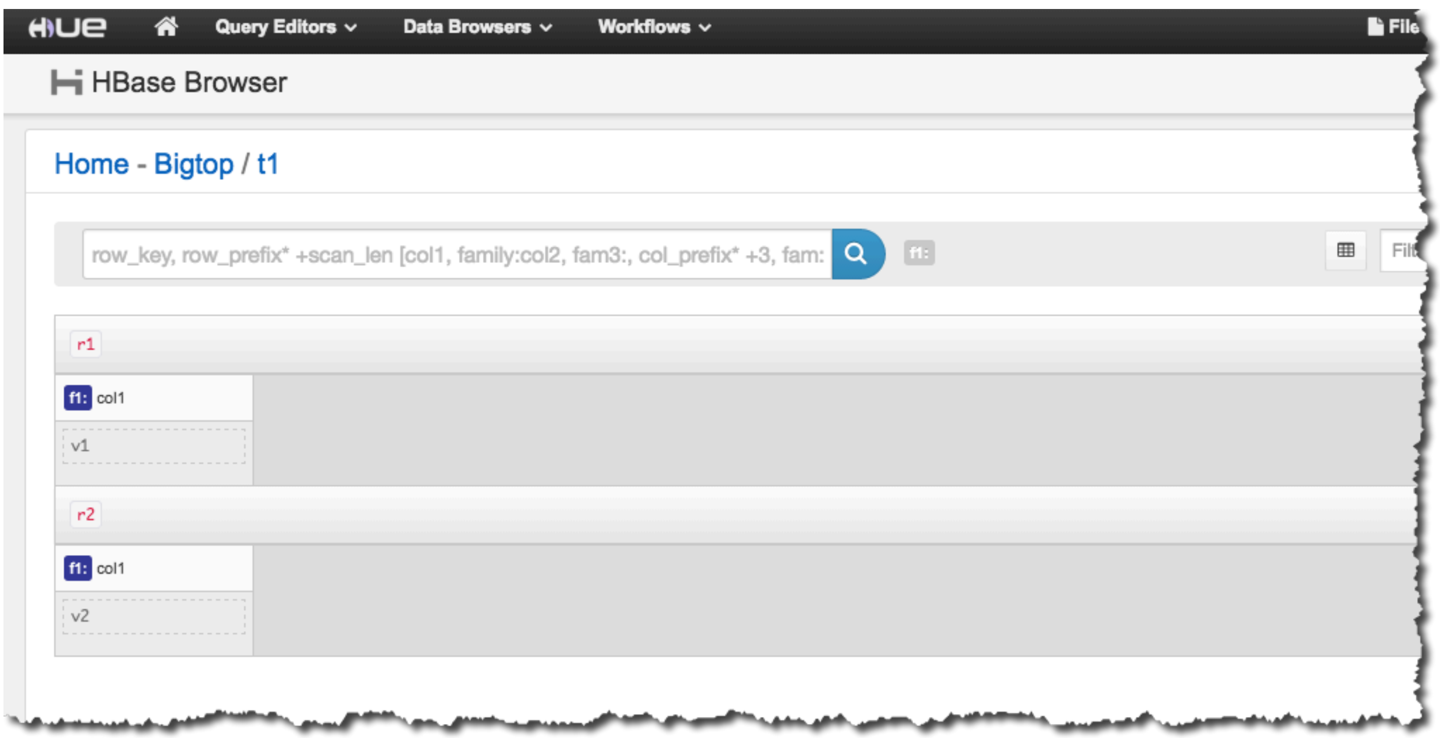
HBase는 HBase 클러스터를 모니터링하는 데 사용할 수 있는 웹 기반 사용자 인터페이스를 제공합니다. Amazon EMR에서 HBase를 실행하면 웹 인터페이스가 프라이머리 노드에서 실행되며 SSH 터널 생성이라고도 하는 포트 전달을 사용하여 볼 수 있습니다.

HBase 사용자 인터페이스를 보는 방법

1. SSH를 사용하여 프라이머리 노드로 터널링하고 보안 연결을 생성합니다. 자세한 내용은 Amazon EMR 관리 안내서의 옵션 2, 1부: 동적 포트 전달을 사용하여 프라이머리 노드에 대해 SSH 터널 설정을 참조하세요.
2. Firefox용 FoxyProxy 플러그인과 같은 프록시 도구를 사용하여 웹 브라우저를 설치하여 AWS 도메인에 대한 SOCKS 프록시를 생성합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 옵션 2, 2부: 프라이머리 노드에 호스팅된 웹 사이트를 볼 수 있도록 프록시 설정 구성을 참조하세요.
3. 프록시 세트와 SSH 연결이 열린 상태에서 http://***master-public-dns-name***:16010/master-status가 포함된 브라우저 창을 열어서 HBase UI를 볼 수 있습니다. 여기서, *master-public-dns-name*은 클러스터 프라이머리 노드의 퍼블릭 DNS 주소입니다.



Hue에서 HBase를 볼 수도 있습니다. 예를 들어, 다음에서는 t1에서 생성된 테이블 HBase 쉘 사용을 보여줍니다.



Hue에 대한 자세한 내용은 Hue 단원을 참조하십시오.

HBase 로그 파일 보기

작업 중에 HBase는 구성 설정, 데몬 작업 및 예외에 대한 세부 정보가 포함된 로그 파일을 작성합니다. 이러한 로그 파일은 HBase의 문제를 디버깅하고 성능을 추적하는 데 유용할 수 있습니다.

로그 파일을 Amazon S3에 유지하도록 클러스터를 구성하는 경우, 로그가 5분마다 Amazon S3에 기록된다는 점을 알아야 합니다. 따라서 최신 로그 파일을 사용할 수 있으려면 약간의 지연이 있을 수 있습니다.

프라이머리 노드에서 HBase 로그를 보는 방법

- SSH를 사용하여 프라이머리 노드에 연결하고 `/var/log/hbase` 디렉터리로 이동하여 현재 HBase 로그를 볼 수 있습니다. 클러스터가 시작될 때 Amazon S3에 대한 로깅이 활성화되지 않은 경우 이 로그는 클러스터가 종료된 후 사용할 수 없습니다.

Amazon S3에서 HBase 로그를 보는 방법

- Amazon S3에서 HBase 로그 및 기타 클러스터 로그에 액세스하고 클러스터가 종료된 후에 사용할 수 있게 하려면 클러스터를 생성할 때 Amazon S3 버킷이 이러한 로그를 수신하도록 지정합니다. `--log-uri` 옵션을 사용하여 이를 수행합니다. 클러스터의 로깅을 활성화하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 로깅 및 디버깅 구성(선택 사항)을 참조하세요.

Ganglia를 사용하여 HBase 모니터링

> **Note**
>
> Ganglia를 포함하는 Amazon EMR의 마지막 릴리스는 Amazon EMR 6.15.0이었습니다. 클러스터를 모니터링하기 위해 6.15.0 이상의 릴리스에는 Amazon CloudWatch 에이전트가 포함됩니다.

Ganglia 오픈 소스 프로젝트는 성능에 미치는 영향을 최소화하면서 클러스터 및 그리드를 모니터링하도록 설계된 확장 가능한 분산 시스템입니다. 클러스터에서 Ganglia를 활성화하면 보고서를 생성하고 전체 클러스터의 성능을 볼 수 있으며 개별 노드 인스턴스의 성능을 검사할 수 있습니다. Ganglia 오픈 소스 프로젝트에 대한 자세한 내용은 http://ganglia.info/를 참조하십시오. Amazon EMR 클러스터와 함께 Ganglia를 사용하는 데 대한 자세한 내용은 Ganglia 섹션을 참조하세요.

Ganglia가 구성된 클러스터가 시작되면 프라이머리 노드에서 실행중인 그래픽 인터페이스를 사용하여 Ganglia 그래프와 보고서에 액세스할 수 있습니다.

Ganglia는 `/mnt/var/lib/ganglia/rrds/` 디렉터리의 프라이머리 노드에 로그 파일을 저장합니다. 이전 릴리스 버전의 Amazon EMR에서는 `/var/log/ganglia/rrds/` 디렉터리에 로그 파일을 저장할 수도 있습니다.

를 사용하여 Ganglia 및 HBase에 대한 클러스터를 구성하려면 AWS CLI

- 다음과 유사한 `create-cluster` 명령을 사용합니다.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=HBase Name=Ganglia --use-default-roles \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3
```

> **Note**
> 기본 Amazon EMR 서비스 역할과 Amazon EC2 인스턴스 프로파일이 없는 경우 오류가 발생합니다. `aws emr create-default-roles` 명령을 사용하여 해당 역할 및 프로파일을 생성한 다음 다시 시도하십시오.

자세한 내용은 [AWS CLI에서 Amazon EMR 명령]을 참조하세요.

Ganglia 웹 인터페이스에서 HBase 지표를 보려면

1. SSH를 사용하여 프라이머리 노드로 터널링하고 보안 연결을 생성합니다. 자세한 내용은 Amazon EMR 관리 안내서의 [옵션 2, 1부: 동적 포트 전달을 사용하여 프라이머리 노드에 대해 SSH 터널 설정]을 참조하세요.
2. Firefox용 FoxyProxy 플러그인과 같은 프록시 도구를 사용하여 웹 브라우저를 설치하여 AWS 도메인에 대한 SOCKS 프록시를 생성합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [옵션 2, 2부: 프라이머리 노드에 호스팅된 웹 사이트를 볼 수 있도록 프록시 설정 구성]을 참조하세요.
3. 프록시 세트와 SSH 연결이 열린 상태에서 http://*master-public-dns-name*/ganglia/가 포함된 브라우저 창을 열어서 Ganglia 지표를 볼 수 있습니다. 여기서 `master-public-dns-name`은 HBase 클러스터의 마스터 서버의 퍼블릭 DNS 주소입니다.

프라이머리 노드에서 Ganglia 로그 파일을 보는 방법

- 클러스터가 아직 실행 중인 경우 SSH를 사용하여 프라이머리 노드에 연결하고 `/mnt/var/lib/ganglia/rrds/` 디렉터리로 이동하여 로그 파일에 액세스할 수 있습니다. EMR 3.x에서는 `/var/log/ganglia/rrds` 디렉터리로 이동합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [SSH를 사용하여 프라이머리 노드에 연결]을 참조하세요.

Amazon S3에서 Ganglia 로그 파일을 보는 방법

- Ganglia 로그 파일은 클러스터에 대해 로깅을 활성화하더라도 Amazon S3에 자동으로 기록되지 않습니다. Amazon S3에서 Ganglia 로그 파일을 보려면 수동으로 `/mnt/var/lib/ganglia/rrds/`에서 S3 버킷으로 로그를 푸시해야 합니다.

Amazon CloudWatch를 사용하여 EMR HBase 모니터링

EMR 7.0부터 Amazon EMR은 지표를 Amazon CloudWatch 또는 Prometheus로 전송할 Amazon CloudWatch 에이전트를 제공하여 Ganglia 모니터링 시스템을 대체합니다. 자세한 내용은 Amazon CloudWatch 에이전트 구성 안내서에서 확인할 수 있습니다.

EMR 7.0 Amazon CloudWatch 에이전트는 기본 통합을 제공하고 구성 요소 다운로드 및 부트스트랩 작업 사용 등 필요한 수동 작업을 설정했습니다. 그러나 EMR 7.1 이상에서는 프로세스가 간소화됩니다. EMR 7.1에서 Amazon CloudWatch Agent를 사용하면 HBase에 대한 지표를 포함하여 시스템 수준 및 애플리케이션별 지표를 모두 쉽게 모니터링할 수 있습니다. EMR 구성 API를 사용하면 지표 수집 프로세스를 빠르게 설정 및 사용자 지정하고 Amazon CloudWatch 또는 Prometheus로 데이터를 전송할 위치를 선택할 수 있습니다. 이러한 유연성은 HBase 클러스터를 면밀히 관찰하여 원활하고 효율적으로 실행되도록 하는 데 도움이 됩니다.

EMR 구성 API를 사용하여 전체 설정을 처리하면 프로세스가 훨씬 원활해집니다. EMR 7.1의 CloudWatch 에이전트는 세 가지 주요 유형의 지표를 지원합니다.

- 시스템 지표 - 여기에는 CPU 사용량, 디스크 사용량, 메모리 사용량, 네트워크 I/O, 프로세스 및 스왑 사용량과 같은 시스템 성능의 주요 지표가 포함됩니다.
- 하둡 데몬 지표 - 이러한 지표는 DataNode 지표, NameNode 지표, YARN NodeManager 지표, YARN ResourceManager 지표를 비롯한 하둡의 다양한 구성 요소와 관련이 있습니다.
- HBase 지표 - 이러한 지표는 HBase 마스터 지표, HBase HBase 리전 서버 지표, HBase REST 서버 지표, HBase Thrift Server 지표 등 HBase의 성능에 대한 인사이트를 제공합니다.

Using the AWS CLI

하둡 및 HBase에 대한 모든 지표는 JMX 기반이므로 Java Management Extensions를 사용하여 자세한 인사이트를 제공합니다. HBase를 모니터링하도록 Amazon CloudWatch Agent를 설정하는 방법은 다음과 같습니다.

- Amazon CloudWatch Agent로 클러스터를 생성하기 전에 사전 조건을 참조하세요. 이 목록 뒤에 나타나는 샘플과 유사한 `create-cluster` 명령을 사용합니다.

- HBase 클러스터에 지원되는 구성을 참조하세요.
- HBase 모니터링을 설정하는 구성 예제는 다음 샘플을 참조하세요. `--configuration` 입력에 대한 구성 예제를 참조하세요.

```
aws emr create-cluster --name "HBase cluster with CloudWatch agent" \
--release-label emr-7.1.0 \
--applications Name=HBase Name=AmazonCloudWatchAgent \
--ec2-attributes KeyName=myKey --instance-type m7g.2xlarge \
--configurations file://./configurations.json \
--instance-count 3 --use-default-roles
```

지표의 JSON 샘플 구성에 대한 자세한 내용은 지표 설정을 참조하세요.

Using the console

콘솔에서 Amazon CloudWatch 에이전트로 클러스터를 생성하려면 다음 단계를 수행합니다.

콘솔에서 CloudWatch 에이전트를 사용하여 클러스터 생성

1. 에 로그인 AWS Management Console하고 https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 생성을 선택합니다.
3. 이름 및 애플리케이션에서 Amazon EMR 릴리스 7.0.0 이상을 선택합니다.
4. 애플리케이션 번들에서 클러스터에 설치할 HBase 및 추가 앱을 선택하고 선택한 항목에 CloudWatch 에이전트를 포함합니다.
5. 소프트웨어 설정을 확장합니다. 그런 다음 JSON 또는 콘솔에 섀도우 텍스트로 표시되는 간편 구문을 사용하여 구성을 직접 입력할 수 있습니다. 그렇지 않으면 파일의 Amazon S3 URI를 JSON `Configurations` 객체와 함께 제공할 수 있습니다. 지표의 JSON 샘플 구성에 대한 자세한 내용은 지표 설정을 참조하세요.
6. 사용 사례 요구 사항에 맞게 클러스터를 생성하세요.

지표 설정

HBase 마스터를 모니터링하려면 특정 지표를 수집하도록 Amazon CloudWatch 에이전트를 설정할 수 있습니다.

1. HBase 마스터 지표 설정 - HBase 마스터를 모니터링하려면 특정 지표를 수집하도록 Amazon CloudWatch Agent를 설정할 수 있습니다. 다음은 마스터 할당 관리자 활동을 추적하는 구성 예제입니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-master-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Master,sub=AssignmentManager":
 "AssignFailedCount,AssignSubmittedCount",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

   이 설정에서 다음을 수행합니다.

   - `AssignFailedCount` 및와 같은 지표를 수집하기 위해 MBean(`Hadoop:service=HBase,name=Master,sub=AssignmentManager`)을 지정합니다`AssignSubmittedCount`.
   - 30초(30,000밀리초)마다 이러한 지표를 수집하도록 간격을 설정합니다.

2. HBase 리전 서버 지표 설정 - HBase 리전 서버를 모니터링하려면 다음과 같이 CloudWatch 에이전트를 구성합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-region-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=RegionServer,sub=IPC":
 "numActiveHandler,numActivePriorityHandler",
```

```
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

이 구성은 다음과 같습니다.

- 리전 서버(`numActiveHandler`, `numActivePriorityHandler`)에서 활성 핸들러를 모니터링합니다.
- 지표 수집에 30초 간격을 사용합니다.

3. HBase REST 서버 지표 설정 - HBase REST 인터페이스를 모니터링하려면 다음 구성을 사용할 수 있습니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-rest-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=REST": "successfulPut,successfulScanCount",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

이 예제에서 CloudWatch Agent는 성공한 PUT 작업에 대한 지표를 수집하고 30초마다 카운트를 스캔합니다.

4. HBase Thrift Server 지표 설정 - HBase Thrift Server를 모니터링하려면 다음과 같은 구성으로 지표를 구성할 수 있습니다.

```
[
  {
```

```
    "Classification": "emr-metrics",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "emr-hbase-thrift-server-metrics",
        "Properties": {
          "Hadoop:service=HBase,name=Thrift,sub=ThriftOne":
 "BatchGet_max,BatchGet_mean",
          "otel.metric.export.interval": "30000"
        },
        "Configurations": []
      }
    ]
  }
]
```

이 설정은 Thrift 서버에서 배치 GET 작업의 최대 및 평균 시간을 추적합니다.

지표 대상 사용

EMR 7.1을 사용하면 지표 데이터를 Amazon CloudWatch 또는 Amazon Managed Service for Prometheus로 전송할 수 있습니다. 이 옵션을 사용하면 필요에 따라 다양한 모니터링 도구와 원활하게 통합할 수 있습니다.

Amazon CloudWatch로 지표 전송

CloudWatch로 지표를 보내려면 다음 구성을 사용합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "cloudwatch"
    },
    "Configurations": []
  }
]
```

Amazon Managed Service for Prometheus로 지표 전송

Prometheus를 사용하려는 경우 대상을 설정하고 엔드포인트 URL을 제공합니다.

```
[
  {
    "Classification": "emr-metrics",
    "Properties": {
      "metrics_destination": "prometheus",
      "prometheus_endpoint": "https://aps-workspaces.region.amazonaws.com/
workspaces/workspace_id/api/v1/remote_write"
    },
    "Configurations": []
  }
]
```

region를 AWS 리전으로 바꾸고를 Prometheus 워크스페이스 IDworkspace_id로 바꿉니다. 이 구성은 지정된 엔드포인트를 사용하여 HBase 지표를 Prometheus로 전달합니다.

위의 설정을 사용하면 모니터링 탭에서 아래 지표를 볼 수 있습니다.

이전 HBase 버전에서 마이그레이션

이전 HBase 버전에서 데이터를 마이그레이션하려면 Apache HBase 참조 설명서의 Upgrading 및HBase version number and compatibility를 참조하십시오. HBase의 1.0 이전 버전에서 업그레이드하기 위한 요구 사항에 특히 주의해야 할 수 있습니다.

Amazon EMR 버전 7.4.0 이상으로 마이그레이션

> Note
>
> 7.4.0 이전의 EMR 릴리스에서 7.3.0 이상의 릴리스로 마이그레이션하는 경우 다음 지침을 따르세요.

현재 버전 6.2.0~7.3.0에 포함된 Amazon의 Store File Tracking 기능이 활성화된 EMR 버전을 실행 중이고 7.3.0 이후 EMR 버전에서 사용할 수 있는 OSS Store File Tracking이 있는 버전으로 업그레이드하려는 경우 다음 단계를 따르세요.

1. 기존 클러스터에서:

    1. hbase:storefile 테이블을 비활성화합니다.
    2. hbase:storefile 테이블을 삭제합니다.

3. 를 플러시합니다`hbase:meta`.
4. 메타데이터가 업데이트될 때까지 기다립니다.

2. 새 클러스터에서:

1. 루트 디렉터리와 동일한 Amazon S3 디렉터리를 설정합니다.
2. `DefaultStoreFileTracker` 구현을 통해 클러스터를 시작합니다.

```
{
  "Classification": "hbase-site",
  "Properties": {
    hbase.store.file-tracker.impl:
 "org.apache.hadoop.hbase.regionserver.storefiletracker.DefaultStoreFileTracker"
}
```

3. 테이블 또는 열 패밀리 수준에서 다음 명령을 사용하여 저장 파일 트래커를 변경합니다.

a. 테이블 또는 테이블 열 패밀리의 Store File Tracker를 변경합니다.

```
hbase> change_sft 't1','FILE'
hbase> change_sft 't2','cf1','FILE'
```

b. 지정된 정규식(정규식)과 일치하는 테이블의 Store File Tracker를 모두 변경합니다.

```
hbase> change_sft_all 't.*','FILE'
hbase> change_sft_all 'ns:.*','FILE'
hbase> change_sft_all 'ns:t.*','FILE'
```

HBase 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 HBase의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

HBase 버전 정보

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.9.0	2.6.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-7.8.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server,

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
		hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-7.7.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.6.0	2.6.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.5.0	2.5.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.4.0	2.5.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.3.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.2.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.1.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-7.0.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.15.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.14.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.13.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.12.0	2.4.17	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.11.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.11.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.10.1	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.10.0	2.4.15	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-wal-cli, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.9.1	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.9.0	2.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.8.1	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-6.8.0	2.4.12	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.7.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server
emr-5.36.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.36.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.6.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, hbase-operator-tools, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.35.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.5.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.4.0	2.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.3.1	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.3.0	2.2.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.2.1	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.2.0	2.2.6-amzn-0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.1.1	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.1.0	2.2.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-6.0.1	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-6.0.0	2.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.34.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.33.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.33.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.32.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.32.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.31.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.31.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.30.2	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.30.1	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.30.0	1.4.13	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.29.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.28.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.28.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.27.1	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.27.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.26.0	1.4.10	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.25.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.24.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.24.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.23.1	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.23.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.22.0	1.4.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.2	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.21.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.21.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.20.1	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.20.0	1.4.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.17.2	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.17.1	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.17.0	1.4.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.16.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.16.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.15.1	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.15.0	1.4.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.2	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.14.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.14.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.13.1	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.13.0	1.4.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.12.3	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.2	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.12.1	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.12.0	1.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.11.4	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.11.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.11.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.11.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.10.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.10.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.9.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.9.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.3	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.8.2	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.8.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.8.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.7.1	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.7.0	1.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.6.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.6.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.4	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.5.3	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.2	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.5.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.5.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.4.1	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.4.0	1.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.3.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.3.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.3.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.3	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.2.2	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.2.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.2.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.1.1	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.1.0	1.2.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.0.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-5.0.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-5.0.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.6	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.9.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.9.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.9.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.9.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.5	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.8.4	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.3	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.8.2	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.8.1	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.8.0	1.2.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.4	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.7.3	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.2	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.7.1	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.7.0	1.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	HBase 버전	HBase와 함께 설치된 구성 요소
emr-4.6.1	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server
emr-4.6.0	1.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, hbase-rest-server, hbase-thrift-server, zookeeper-client, zookeeper-server

Apache HCatalog

HCatalog는 Pig, Spark SQL 및/또는 사용자 지정 MapReduce 애플리케이션 내의 Hive 메타스토어 테이블에 액세스할 수 있는 도구입니다. HCatalog에는 테이블을 생성하거나 기타 작업을 수행할 수 있는 REST 인터페이스 및 명령줄 클라이언트가 있습니다. 그런 다음 애플리케이션을 작성하여 HCatalog 라이브러리를 사용하는 테이블에 액세스합니다. 자세한 내용은 Using HCatalog를 참조하세요. HCatalog는 Amazon EMR 릴리스 버전 4.4.0 이상에 포함되어 있습니다.

Amazon EMR 릴리스 버전 5.8.0 이상의 HCatalog는 Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용할 수 있도록 지원합니다. 자세한 내용은 Glue 데이터 카탈로그를 Hive AWS 의 메타스토어로 사용을 참조하세요.

다음 테이블에는 Amazon EMR이 HCatalog를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 HCatalog의 버전이 나열되어 있습니다.

이 릴리스에서 HCatalog와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 HCatalog 버전 정보

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.15.0	HCatalog 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

다음 테이블에는 Amazon EMR이 HCatalog를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 HCatalog의 버전이 나열되어 있습니다.

이 릴리스에서 HCatalog와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 HCatalog 버전 정보

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.36.2	HCatalog 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

주제

- HCatalog를 포함하는 클러스터 생성
- HCatalog 사용
- 예제: HCatalog 테이블 생성 및 Pig를 사용하여 테이블에 쓰기
- HCatalog 릴리스 기록

HCatalog를 포함하는 클러스터 생성

HCatalog는 Hive 프로젝트에 포함되어 있지만, 독자적인 애플리케이션으로 설치해야 합니다.

콘솔을 사용하여 설치된 HCatalog와 함께 클러스터를 시작하려면

다음 절차에서는 HCatalog가 설치된 클러스터를 생성합니다. 고급 옵션을 비롯해 콘솔을 사용하여 클러스터를 생성하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 계획 및 구성을 참조하세요.

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 빠른 생성을 사용하려면 클러스터 생성을 선택합니다.
3. 소프트웨어 구성 필드에 대해 Amazon 릴리스 버전 emr-4.4.0 이상을 선택합니다.
4. Select Applications(애플리케이션 선택) 필드에서 모든 애플리케이션 또는 HCatalog를 선택합니다.
5. 필요에 따라 다른 옵션을 선택한 다음 클러스터 생성을 선택합니다.

를 사용하여 HCatalog로 클러스터를 시작하려면 AWS CLI

- 다음 명령을 사용하여 클러스터를 생성합니다.

  Note

  가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Cluster with Hcat" --release-label emr-7.9.0 \
--applications Name=HCatalog --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

HCatalog 사용

Hive 메타스토어를 사용하는 다양한 애플리케이션 내에서 HCatalog를 사용할 수 있습니다. 이 단원의 예제에서는 테이블을 생성하고 Pig 및 Spark SQL의 컨텍스트에서 테이블을 사용하는 방법을 보여 줍니다.

HCatalog HStorer 사용 시 직접 쓰기 비활성화

애플리케이션에서 HCatStorer를 사용하여 Amazon S3에 저장된 HCatalog 테이블에 쓸 때마다 Amazon EMR의 직접 쓰기 기능을 비활성화합니다. 예를 들면, Pig STORE 명령

을 사용할 때나 Amazon S3에 HCatalog 테이블을 쓰는 Sqoop 작업을 실행할 때는 직접 쓰기를 비활성화합니다. `mapred.output.direct.NativeS3FileSystem` 및 `mapred.output.direct.EmrFileSystem` 구성을 `false`로 설정하여 직접 쓰기 기능을 비활성화할 수 있습니다. 다음 예제에서는 Java를 사용하여 이러한 구성을 설정하는 방법을 보여 줍니다.

```
Configuration conf = new Configuration();
conf.set("mapred.output.direct.NativeS3FileSystem", "false");
conf.set("mapred.output.direct.EmrFileSystem", "false");
```

HCat CLI를 사용하여 테이블 생성 및 Pig에서 해당 데이터 사용

클러스터에서 다음 `impressions.q` 스크립트를 생성합니다.

```
CREATE EXTERNAL TABLE impressions (
    requestBeginTime string, adId string, impressionId string, referrer string,
    userAgent string, userCookie string, ip string
  )
  PARTITIONED BY (dt string)
  ROW FORMAT
    serde 'org.apache.hive.hcatalog.data.JsonSerDe'
    with serdeproperties ( 'paths'='requestBeginTime, adId, impressionId, referrer,
 userAgent, userCookie, ip' )
  LOCATION 's3://[your region].elasticmapreduce/samples/hive-ads/tables/impressions/';
ALTER TABLE impressions ADD PARTITION (dt='2009-04-13-08-05');
```

HCat CLI를 사용하여 스크립트 실행:

```
% hcat -f impressions.q
Logging initialized using configuration in file:/etc/hive/conf.dist/hive-
log4j.properties
OK
Time taken: 4.001 seconds
OK
Time taken: 0.519 seconds
```

Grunt 셸을 열고 `impressions`의 데이터에 액세스합니다.

```
% pig -useHCatalog -e "A = LOAD 'impressions' USING
 org.apache.hive.hcatalog.pig.HCatLoader();
B = LIMIT A 5;
dump B;"
```

```
<snip>
(1239610346000,m9nwdo67Nx6q2kI25qt5On7peICfUM,omkxkaRpNhGPDucAiBErSh1cs0MThC,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 6.0; FunWebProducts; GTB6; SLCC1; .NET CLR
 2.0.50727; Media Center PC
 5.0; .NET,wcVWWTascoPbGt6bdqDbuWTPPHgOPs,69.191.224.234,2009-04-13-08-05)
(1239611000000,NjriQjdODgWBKnkGJUP6GNTbDeK4An,AWtXPkfaWGOaNeL9OOsFU8Hcj6eLHt,cartoonnetwork.com
 (compatible; MSIE 7.0; Windows NT 5.1; GTB6; .NET CLR
 1.1.4322),OaMU1F2gE4CtADVHAbKjjRRks5kIgg,57.34.133.110,2009-04-13-08-05)
(1239610462000,Irpv3oiu0I5QNQiwSSTIshrLdo9cM1,i1LDq44LRSJF0hbmhB8Gk7k9gMWtBq,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239611007000,q2Awfnpe0JAvhInaIp0VGx9KTs0oPO,s3HvTflPB8JIE0IuM6hOEebWWpOtJV,cartoonnetwork.com
 (compatible; MSIE 6.0; Windows NT 5.2; SV1; .NET CLR 1.1.4322;
 InfoPath.1),QSb3wkLR4JAIut4Uq6FNFQIR1rCVwU,42.174.193.253,2009-04-13-08-05)
(1239610398000,c362vpAB0soPKGHRS43cj6TRwNeOGn,jeas5nXbQInGAgFB8jlkhnprN6cMw7,cartoonnetwork.com
 (compatible; MSIE 8.0; Windows NT 5.1; Trident/4.0; GTB6; .NET CLR
 1.1.4322),k96n5PnUmwHKfiUI0TFP0TNMfADgh9,51.131.29.87,2009-04-13-08-05)
7120 [main] INFO  org.apache.pig.Main  - Pig script completed in 7 seconds and 199
 milliseconds (7199 ms)
16/03/08 23:17:10 INFO pig.Main: Pig script completed in 7 seconds and 199 milliseconds
 (7199 ms)
```

Spark SQL을 사용하여 테이블에 액세스

이 예제에서는 첫 번째 예제에서 생성된 테이블에서 Spark DataFrame을 생성하고 처음 20줄을 보여 줍니다.

```
% spark-shell --jars /usr/lib/hive-hcatalog/share/hcatalog/hive-hcatalog-core-1.0.0-
amzn-3.jar
<snip>
scala> val hiveContext = new org.apache.spark.sql.hive.HiveContext(sc);
scala> val df = hiveContext.sql("SELECT * FROM impressions")
scala> df.show()
<snip>
16/03/09 17:18:46 INFO DAGScheduler: ResultStage 0 (show at <console>:32) finished in
 10.702 s
16/03/09 17:18:46 INFO DAGScheduler: Job 0 finished: show at <console>:32, took
 10.839905 s
+----------------+--------------------+--------------------+------------------
+--------------------+--------------------+--------------+----------------+
|requestbegintime|                adid|        impressionid|          referrer|
   useragent|          usercookie|            ip|              dt|
```

```
+----------------+-------------------+-------------------+------------------
+-------------------+-------------------+-------------+---------------+
|   1239610346000|m9nwdo67Nx6q2kI25...|omkxkaRpNhGPDucAi...|cartoonnetwork.com|
Mozilla/4.0 (comp...|wcVWWTascoPbGt6bd...|69.191.224.234|2009-04-13-08-05|
|   1239611000000|NjriQjdODgWBKnkGJ...|AWtXPkfaWGOaNeL9O...|cartoonnetwork.com|
Mozilla/4.0 (comp...|OaMU1F2gE4CtADVHA...| 57.34.133.110|2009-04-13-08-05|
|   1239610462000|Irpv3oiu0I5QNQiwS...|i1LDq44LRSJF0hbmh...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239611007000|q2Awfnpe0JAvhInaI...|s3HvTflPB8JIE0IuM...|cartoonnetwork.com|
Mozilla/4.0 (comp...|QSb3wkLR4JAIut4Uq...|42.174.193.253|2009-04-13-08-05|
|   1239610398000|c362vpAB0soPKGHRS...|jeas5nXbQInGAgFB8...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610600000|cjBTpruoaiEtqLuMX...|XwlohBSs8Ipxs1bRa...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610804000|Ms3eJHNAEItpxvimd...|4SIj4pGmgVLl625BD...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610872000|h5bccHX6wJReDi1jL...|EFAWIiBdVfnxwAMWP...|cartoonnetwork.com|
Mozilla/4.0 (comp...|k96n5PnUmwHKfiUI0...|  51.131.29.87|2009-04-13-08-05|
|   1239610365000|874NBpGmxNFfxEPKM...|xSvE4XtGbdtXPF2Lb...|cartoonnetwork.com|
Mozilla/5.0 (Maci...|eWDEVVUphlnRa273j...| 22.91.173.232|2009-04-13-08-05|
|   1239610348000|X8gISpUTSqh1A5reS...|TrFblGT99AgE75vuj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610743000|kbKreLWB6QVueFrDm...|kVnxx9Ie2i3OLTxFj...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610812000|9lxOSRpEi3bmEeTCu...|1B2sff99AEIwSuLVV...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610876000|lijjmCf2kuxfBTnjL...|AjvufgUtakUFcsIM9...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610941000|t8t8trgjNRPIlmxuD...|agu2u2TCdqWP08rAA...|       corriere.it|
Mozilla/4.0 (comp...|tX1sMpnhJUhmAF7AS...|   55.35.44.79|2009-04-13-08-05|
|   1239610490000|OGRLPVNGxiGgrCmWL...|mJg2raBUpPrC8OlUm...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610556000|OnJID12x0RXKPUgrD...|P7Pm2mPdW6wO8KA3R...|       corriere.it|
Mozilla/4.0 (comp...|r2k96t1CNjSU9fJKN...|   71.124.66.3|2009-04-13-08-05|
|   1239610373000|WflsvKIgOqfIE5KwR...|TJHd1VBspNcua0XPn...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610768000|4MJR0XxiVCU1ueXKV...|1OhGWmbvKf8ajoU8a...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610832000|gWIrpDiN57i3sHatv...|RNL4C7xPi3tdar2Uc...|       corriere.it|
Mozilla/5.0 (Maci...|fj2L1ILTFGMfhdrt3...| 75.117.56.155|2009-04-13-08-05|
|   1239610789000|pTne9k62kJ14QViXI...|RVxJVIQousjxUVI3r...|        pixnet.net|
Mozilla/5.0 (Maci...|1bGOKiBD2xmui9OkF...| 33.176.101.80|2009-04-13-08-05|
+----------------+-------------------+-------------------+------------------
+-------------------+-------------------+-------------+---------------+
```

```
only showing top 20 rows


scala>
```

예제: HCatalog 테이블 생성 및 Pig를 사용하여 테이블에 쓰기

Amazon S3에서 데이터 소스를 사용하는 HCatStorer를 통해 HCatalog 테이블을 생성하고 Apache Pig를 사용하여 테이블에 쓸 수 있습니다. HCatalog를 사용하려면 직접 쓰기를 비활성화해야 합니다. 그렇지 않으면 작업이 자동으로 실패합니다. `mapred.output.direct.NativeS3FileSystem` 분류를 사용하거나 Grunt 셸에서 수동으로 `mapred.output.direct.EmrFileSystem`및 `false` 구성을 모두 `mapred-site`로 설정합니다. 다음 예제에서는 HCat CLI를 사용하여 생성된 테이블을 보여 준 다음, Amazon S3의 샘플 데이터 파일에서 테이블을 채우기 위해 Grunt 쉘에서 실행되는 명령을 보여 줍니다.

이 예제를 실행하려면 [SSH를 사용하여 프라이머리 노드에 연결](#)합니다.

다음 콘텐츠가 포함된 HCatalog 스크립트 파일 `wikicount.q`를 생성합니다. 이 파일은 `wikicount`라는 이름의 HCatalog 테이블을 생성합니다.

```
CREATE EXTERNAL TABLE IF NOT EXISTS wikicount(
col1 string,
col2 bigint
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\001'
STORED AS ORC
LOCATION 's3://amzn-s3-demo-bucket/hcat/wikicount';
```

HCat CLI 명령을 사용하여 파일에서 스크립트를 실행합니다.

```
hcat -f wikicount.q
```

다음에는 `-useHCatalog` 옵션을 선택하여 Grunt 셸을 시작하고, 직접 쓰기를 비활성화하도록 구성을 설정한 다음, S3 위치에서 데이터를 로드하고, 결과를 wikicount 테이블에 기록합니다.

```
pig -useHCatalog
SET mapred.output.direct.NativeS3FileSystem false;
```

```
SET mapred.output.direct.EmrFileSystem false;
A = LOAD 's3://support.elasticmapreduce/training/datasets/wikistats_tiny/' USING
 PigStorage(' ') AS (Site:chararray, page:chararray, views:int, total_bytes:long);
B = GROUP A BY Site;
C = FOREACH B GENERATE group as col1, COUNT(A) as col2;
STORE C INTO 'wikicount' USING org.apache.hive.hcatalog.pig.HCatStorer();
```

HCatalog 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 HCatalog의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

HCatalog 버전 정보

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
		yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-7.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-5.36.2	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-5.36.1	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.36.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.35.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mariadb-s erver

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mariadb-server
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mysql-server
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mysql-server
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mysql-server
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hcatalog-client, hcatalog-server, hcatalog-webhcat-s erver, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Amazon EMR 릴리스 레이블	HCatalog 버전	HCatalog와 함께 설치된 구성 요소
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hcatalog-client, hcatalog-server, hcatalog-webhcat-server, hive-client, hive-metastore-server, mysql-server

Apache Hive

Hive는 하둡 클러스터를 기반으로 실행되는 오픈 소스 데이터 웨어하우스 및 분석 패키지입니다. Hive 스크립트는 프로그래밍 모델을 추상화하고 일반적인 데이터 웨어하우스 상호 작용을 지원하는 Hive QL(쿼리 언어)이라는 SQL 유형의 언어를 사용합니다. Hive를 사용하면 Java와 같은 더 낮은 수준의 컴퓨터 언어로 MapReduce 프로그램 또는 DAG(방향성 비순환 그래프) 기반 Tez 작업을 작성하는 복잡성을 피할 수 있습니다.

Hive에서는 직렬화 형식을 포함하여 SQL 패러다임을 확장합니다. 데이터 자체를 건드리지 않고 데이터와 일치하는 테이블 스키마를 생성하여 쿼리 처리를 사용자 지정할 수도 있습니다. SQL은 기본적인 값 유형(예: 날짜, 숫자 및 문자열)만 지원하는 반면, Hive 테이블 값은 JSON 객체, 사용자 정의 데이터 유형, Java로 작성된 함수 등과 같은 구조적 요소입니다.

Hive에 대한 자세한 내용은 http://hive.apache.org/를 참조하십시오.

다음 테이블에는 Amazon EMR이 Hive를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Hive의 버전이 나열되어 있습니다.

이 릴리스에서 Hive와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Hive 버전 정보

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-7.9.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
		worker, zookeeper-client, zookeeper-server

다음 테이블에는 Amazon EMR이 Hive를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Hive의 버전이 나열되어 있습니다.

이 릴리스에서 Hive와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Hive 버전 정보

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.15.0	Hive 3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

다음 테이블에는 Amazon EMR이 Hive를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Hive의 버전이 나열되어 있습니다.

이 릴리스에서 Hive와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Hive 버전 정보

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.36.2	Hive 2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 5.18.0부터 Amazon EMR 아티팩트 리포지토리를 사용하여 특정 Amazon EMR 릴리스 버전에서 사용할 수 있는 정확한 버전의 라이브러리와 종속 항목에 대해 작업 코드를 빌드할 수 있습니다. 자세한 내용은 Amazon EMR 아티팩트 리포지토리를 사용하여 종속성 확인 단원을 참조하십시오.

주제

- Amazon EMR에서 Hive의 차이점 및 고려 사항
- Hive용 외부 메타스토어 구성
- Hive JDBC 드라이버 사용
- Hive 성능 개선
- Hive Live Long and Process(LLAP) 사용
- Hive에서 암호화
- Hive 릴리스 기록

Amazon EMR에서 Hive의 차이점 및 고려 사항

Amazon EMR의 Apache Hive와 Apache Hive 간의 차이점

이 섹션에서는 Amazon EMR의 Hive와 http://svn.apache.org/viewvc/hive/branches/에서 사용할 수 있는 기본 Hive 버전 간의 차이점을 설명합니다.

Hive 권한 부여

Amazon EMR은 HDFS에 대해 Hive 권한 부여를 지원하지만 EMRFS 및 Amazon S3에서는 지원하지 않습니다. Amazon EMR 클러스터는 기본적으로 권한 부여가 비활성화된 상태로 실행됩니다.

Amazon S3와 Hive 파일 병합 동작

Apache Hive는 `hive.merge.mapfiles`가 true이고 작업의 평균 출력 크기가 `hive.merge.smallfiles.avgsize` 설정보다 작은 경우에만 병합이 트리거되면 맵 전용 작업의 끝에서 작은 파일을 병합합니다. Amazon EMR Hive는 최종 출력 경로가 HDFS에 있는 경우 정확히 동일한 동작을 수행합니다. 출력 경로가 Amazon S3에 있는 경우 `hive.merge.smallfiles.avgsize` 파라미터는 무시됩니다. 이 상황에서 `hive.merge.mapfiles`가 `true`로 설정된 경우 병합 작업이 항상 트리거됩니다.

ACID 트랜잭션 및 Amazon S3

Amazon EMR 6.1.0 이상에서는 Hive 원자성, 일관성, 격리성, 내구성(ACID) 트랜잭션을 지원하므로 데이터베이스의 ACID 속성을 준수합니다. 이 기능을 사용하면 Amazon Simple Storage Service(S3)의 데이터를 사용하여 Hive 관리형 테이블에서 INSERT, UPDATE, DELETE, MERGE 작업을 실행할 수 있습니다.

Hive LLAP(Live Long and Process)

기본 Apache Hive 버전 2.0에서 추가된 LLAP 기능은 Amazon EMR 릴리스 5.0의 Hive 2.1.0에서는 지원되지 않습니다.

Amazon EMR 버전 6.0.0 이상에서는 Hive용 LLAP(Live Long and Process) 기능을 지원합니다. 자세한 내용은 Hive LLAP 사용을 참조하십시오.

Amazon EMR 릴리스 버전 4.x 및 5.x 간 Hive의 차이

이 섹션에서는 Amazon EMR 릴리스 4.x의 Hive 버전 1.0.0에서 Amazon EMR 릴리스 5.x의 Hive 2.x로 Hive 구현을 마이그레이션하기 전에 고려해야 할 차이점을 설명합니다.

작업 차이점 및 고려 사항

- 원자성, 일관성, 격리성, 내구성(ACID) 트랜잭션에 대한 지원이 추가됨: Amazon EMR 4.x의 Hive 1.0.0 및 기본 Apache Hive 간의 이 차이점이 제거되었습니다.
- Amazon S3에 대한 직접 쓰기가 제거됨: Amazon EMR의 Hive 1.0.0과 기본 Apache Hive 간의 이 차이점이 제거되었습니다. 이제 Amazon EMR 릴리스 5.x의 Hive 2.1.0이 Amazon S3에 저장된 임시 파일을 생성하고, 해당 파일에서 읽고 해당 파일에 씁니다. 결과적으로 같은 테이블에서 읽고 쓰기 위해 차선책으로 임시 테이블을 클러스터의 로컬 HDFS 파일 시스템에 만들 필요가 더 이상 없습니다. 버전이 지정된 버킷을 사용하는 경우 아래 설명한 대로 이러한 임시 파일을 관리해야 합니다.
- Amazon S3 버전 지정된 버킷을 사용하는 경우 임시 파일 관리: 생성된 데이터의 대상이 Amazon S3인 Hive 쿼리를 실행하면 많은 임시 파일과 디렉터리가 생성됩니다. 앞에서 설명한대로 새로운 동작입니다. 버전 지정된 S3 버킷을 사용하는 경우 이러한 임시 파일은 Amazon S3를 복잡하게 만들며, 이를 삭제하지 않으면 비용이 발생합니다. 짧은 기간(예: 5일) 후에 `/_tmp` 접두사가 있는 데이터가 삭제되도록 수명 주기 규칙을 조정하십시오. 자세한 내용은 수명 주기 구성 지정을 참조하세요.
- Log4j가 log4j 2로 업데이트됨: log4j를 사용하는 경우 이 업그레이드로 인해 로깅 구성을 변경해야 할 수 있습니다. 자세한 내용은 Apache log4j 2를 참조하십시오.

성능 차이점 및 고려사항

- Tez의 성능 차이점: Amazon EMR 릴리스 5.x에서 Tez는 MapReduce 대신 Hive의 기본 실행 엔진입니다. Tez은 대부분의 워크플로에서 개선된 성능을 제공합니다.
- 파티션이 많은 테이블: 많은 수의 동적 파티션을 생성하는 쿼리는 실패할 수 있으며 많은 파티션이 있는 테이블에서 선택하는 쿼리는 실행하는 데 예상보다 오래 걸릴 수 있습니다. 예를 들어, 100,000개의 파티션 중에서 선택하려면 10분 이상 걸릴 수 있습니다.

Amazon EMR에서 Hive의 추가 기능

Amazon EMR은 Amazon Simple Storage Service(Amazon S3) 및 DynamoDB에서 읽고 쓸 수 있는 기능과 같은 다른 AWS 서비스와 Hive 통합을 지원하는 새로운 기능으로 Hive를 확장합니다.

Hive의 변수

달러 기호와 중괄호를 사용하여 스크립트에 변수를 포함할 수 있습니다.

```
add jar ${LIB}/jsonserde.jar
```

다음 예제와 같이 -d 파라미터를 사용하여 명령줄에서 이러한 변수의 값을 Hive로 전달할 수 있습니다.

```
-d LIB=s3://elasticmapreduce/samples/hive-ads/lib
```

또한 Hive 스크립트를 실행하는 단계로 값을 전달할 수 있습니다.

콘솔을 사용하여 변수 값을 Hive 단계로 전달하려면

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 생성을 선택합니다.
3. Steps(단계) 섹션의 Add Step(단계 추가)에 있는 목록에서 Hive Program(Hive 프로그램)을 선택하고 Configure and add(구성 및 추가)를 선택합니다.
4. Add Step(단계 추가) 대화 상자에서 다음 표를 가이드로 사용하여 파라미터를 지정한 다음 Add(추가)를 선택합니다.

필드	작업
Script S3 location*	스크립트가 위치한 Amazon S3의 URI를 지정합니다. 값은 *BucketName* /*path*/*ScriptName* 형식이어야 합니다. 예를 들어 `s3://elasticmapreduce/samples/hive-ads/libs/response-time-stats.q` 입니다.
Input S3 location	선택적으로 입력 파일이 위치한 Amazon S3의 URI를 지정합니다. 값은 *BucketName* /*path*/ 형식이어야 합니다. 지정된 경우 이 값은 Hive 스크립트에 INPUT 이름의 파라미터로 전달됩니다. 예를 들어 `s3://elasticmapreduce/samples/hive-ads/tables/` 입니다.
Output S3 location	선택적으로 출력을 저장할 Amazon S3의 URI를 지정합니다. 값은 *BucketName* /*path* 형식이어야 합니다. 지정된 경우 이 값은 Hive 스크립트에 OUTPUT 이름의 파라미터로 전달됩니다. 예를 들어 `s3://amzn-s3-demo-bucket/hive-ads/output/` 입니다.

<table>
<tr><th>필드</th><th>작업</th></tr>
<tr><td>인수</td><td>또는 Hive에 전달할 인수 목록(공백으로 구분된 문자열)을 입력합니다. ${SAMPLE} 이름의 경로 변수를 Hive 스크립트에 정의한 경우, 예를 들면 다음과 같습니다.
<pre>
CREATE EXTERNAL TABLE logs (requestBeginTime STRING, requestEn
dTime STRING, hostname STRING) PARTITIONED BY (dt STRING) \
ROW FORMAT serde 'com.amazon.elasticmapreduce.JsonSerde'
WITH SERDEPROPERTIES ('paths'='requestBeginTime, requestEn
dTime, hostname') LOCATION '${SAMPLE}/tables/impressions';
</pre>
변수 값을 전달하려면 Arguments(인수) 창에 다음을 입력합니다.

<code>-d SAMPLE=s3://elasticmapreduce/samples/hive-ads/</code> .</td></tr>
<tr><td>Action on Failure</td><td>이 옵션은 오류 발생 시 클러스터가 수행해야 할 동작을 결정합니다. 이 설정에 사용할 수 있는 값은 다음과 같습니다.

클러스터 종료: 단계가 실패할 경우 클러스터를 종료합니다. 클러스터에 종료 방지 기능과 연결 유지 기능이 활성화된 경우에는 클러스터가 종료되지 않습니다.
취소 및 대기: 단계가 실패할 경우 남은 단계를 취소합니다. 클러스터에 연결 유지 기능이 활성화된 경우에는 클러스터가 종료되지 않습니다.
계속: 단계가 실패할 경우 다음 단계가 계속 수행됩니다.
</td></tr>
</table>

5. 필요에 따라 값을 선택하고 Create cluster(클러스터 생성)를 선택합니다.

를 사용하여 변수 값을 Hive 단계에 전달하려면 AWS CLI

를 사용하여 변수 값을 Hive 단계에 전달하려면 `--steps` 파라미터를 AWS CLI사용하고 인수 목록을 포함합니다.

- > **Note**
  >
  > 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.9.0 \
--applications Name=Hive Name=Pig --use-default-roles --ec2-attributes
 KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Hive,Name="Hive Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/hive-ads/libs/response-time-stats.q,-d,INPUT=s3://
elasticmapreduce/samples/hive-ads/tables,-d,OUTPUT=s3://amzn-s3-demo-bucket/hive-
ads/output/,-d,SAMPLE=s3://elasticmapreduce/samples/hive-ads/]
```

에서 Amazon EMR 명령을 사용하는 방법에 대한 자세한 내용은 섹션을 AWS CLI참조하세요https://docs.aws.amazon.com/cli/latest/reference/emr.

Java SDK;를 사용하여 변수 값을 Hive 단계로 전달하려면

- 다음 예에서는 SDK를 사용하여 변수를 단계에 전달하는 방법을 보여 줍니다. 자세한 내용은 AWS SDK for Java API 참조에서 Class StepFactory를 참조하세요.

```
StepFactory stepFactory = new StepFactory();

   StepConfig runHive = new StepConfig()
     .withName("Run Hive Script")
     .withActionOnFailure("TERMINATE_JOB_FLOW")
     .withHadoopJarStep(stepFactory.newRunHiveScriptStep(“s3://amzn-s3-demo-bucket/
script.q”,
      Lists.newArrayList(“-d”,”LIB= s3://elasticmapreduce/samples/hive-ads/lib”));
```

부분 DynamoDB 스키마를 수용할 Amazon EMR Hive 쿼리

Amazon EMR Hive는 쿼리에 모든 열을 포함하지 않아도 데이터를 필터링할 수 있는 열 하위 세트를 지정할 수 있도록 함으로써 DynamoDB 테이블 쿼리 시 최대 유연성을 제공합니다. 이 부분 스키마 쿼리 기술은 스파스 데이터베이스 스키마가 있고 타임스탬프 필터링 같은 몇 개 열에 기초하여 레코드를 필터링해야 하는 경우에 효과적입니다.

다음 예제에서는 Hive 쿼리를 사용하여 다음을 수행하는 방법을 보여 줍니다.

- DynamoDB 테이블을 생성합니다.
- DynamoDB에서 항목 하위 세트(행)를 선택하고 데이터의 범위를 특정 열로 좁힙니다.
- 결과 데이터를 Amazon S3에 복사합니다.

```
DROP TABLE dynamodb;
DROP TABLE s3;

CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, fullColumn
 map<String, String>)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");

CREATE EXTERNAL TABLE s3(map<String, String>)
     ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
     LOCATION 's3://bucketname/path/subpath/';

INSERT OVERWRITE TABLE s3 SELECT item fullColumn FROM dynamodb WHERE recordTimeStamp <
 "2012-01-01";
```

다음 테이블에는 DynamoDB에서 임의의 항목 조합을 선택하기 위한 쿼리 구문이 표시됩니다.

쿼리 예제	결과 설명
SELECT * FROM *table_name* ;	해당 테이블에서 모든 항목(행)을 선택하고 이러한 항목에 사용할 수 있는 모든 열의 데이터를 포함합니다.
SELECT * FROM *table_name* WHERE *field_name* =*value*;	해당 테이블에서 일부 항목(행)을 선택하고 이러한 항목에 사용할 수 있는 모든 열의 데이터를 포함합니다.
SELECT *column1_name* , *column2_name* , *column3_name* FROM *table_name* ;	해당 테이블에서 모든 항목(행)을 선택하고 이러한 항목에 사용할 수 있는 일부 열의 데이터를 포함합니다.
SELECT *column1_name* , *column2_name* , *column3_name* FROM *table_name* WHERE *field_name* =*value*;	해당 테이블에서 일부 항목(행)을 선택하고 이러한 항목에 사용할 수 있는 일부 열의 데이터를 포함합니다.

다양한 AWS 리전의 DynamoDB 테이블 간에 데이터 복사

Amazon EMR Hive는 DynamoDB 테이블마다 설정할 수 있는 `dynamodb.region` 속성을 제공합니다. `dynamodb.region`을 두 테이블에서 서로 다르게 설정할 경우 테이블 간에 복사하는 모든 데이터가 지정된 리전 사이에서 자동으로 발생합니다.

다음 예제에서는 `dynamodb.region` 속성을 설정하는 Hive 스크립트를 사용하여 DynamoDB 테이블을 생성하는 방법을 보여줍니다.

> **Note**
>
> 테이블당 리전 속성은 전역 Hive 속성을 재정의합니다.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.region" = "eu-west-1",
     "dynamodb.throughput.read.percent" = ".1000",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

테이블당 DynamoDB 처리량 값 설정

Amazon EMR Hive를 사용하면 테이블 정의에서 테이블 단위로 DynamoDB readThroughputPercent 및 writeThroughputPercent 설정을 지정할 수 있습니다. 다음 Amazon EMR Hive 스크립트는 처리량 값을 설정하는 방법을 보여줍니다. DynamoDB 처리량 값에 대한 자세한 내용은 [테이블의 읽기/쓰기 요건 지정]을 참조하세요.

```
CREATE EXTERNAL TABLE dynamodb(hashKey STRING, recordTimeStamp BIGINT, map<String,
 String> fullColumn)
    STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
    TBLPROPERTIES (
     "dynamodb.table.name" = "myTable",
     "dynamodb.throughput.read.percent" = ".4",
     "dynamodb.throughput.write.percent" = "1.0",
     "dynamodb.column.mapping" = "hashKey:HashKey,recordTimeStamp:RangeKey");
```

Hive용 외부 메타스토어 구성

Hive는 기본적으로 프라이머리 노드의 파일 시스템에 위치한 MySQL 데이터베이스에 메타스토어 정보를 기록합니다. 메타스토어에는 파티션 이름, 데이터 유형 등을 포함하여 테이블과 테이블의 기반이 되는 기본 데이터에 대한 설명이 포함되어 있습니다. 클러스터가 종료되면 프라이머리 노드를 포함하여 모든 클러스터 노드가 종료됩니다. 이 경우 노드 파일 시스템에서 휘발성 스토리지를 사용하므로 로컬 데이터가 손실됩니다. 메타스토어를 유지해야 하는 경우 클러스터 외부에 존재하는 외부 메타스토어를 생성해야 합니다.

외부 메타스토어에 대한 두 가지 옵션이 있습니다.

- AWS Glue Data Catalog(Amazon EMR 릴리스 5.8.0 이상만 해당).

  자세한 내용은 Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용 단원을 참조하십시오.
- Amazon RDS 또는 Amazon Aurora.

  자세한 내용은 외부 MySQL 데이터베이스 또는 Amazon Aurora 사용 단원을 참조하십시오.

> **Note**
>
> Hive 3을 사용 중인데 Hive 메타스토어에 너무 많이 연결되면 `datanucleus.connectionPool.maxPoolSize` 파라미터를 더 작은 값으로 구성하거나 데이터베이스 서버가 처리할 수 있는 연결 수를 늘립니다. 연결 수가 증가한 이유는 Hive가 최대 JDBC 연결 수를 계산하는 방식 때문입니다. 최적의 성능 값을 계산하려면 Hive Configuration Properties를 참조하세요.

Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용

Amazon EMR 릴리스 5.8.0 이상을 사용하면 AWS Glue 데이터 카탈로그를 메타스토어로 사용하도록 Hive를 구성할 수 있습니다. 영구 메타스토어가 필요하거나 여러 클러스터, 서비스, 애플리케이션 또는 AWS 계정에서 메타스토어를 공유해야 하는 경우에 이 구성을 사용하는 것이 좋습니다.

AWS Glue는 완전 관리형 ETL(추출, 변환 및 로드) 서비스로, 데이터를 분류하고 정리하고 보강하고 다양한 데이터 스토어 간에 안정적으로 이동할 수 있는 간단하고 비용 효율적입니다. Glue 데이터 카탈로그는 Amazon EMR뿐만 아니라 Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena 및 Apache Hive AWS 메타스토어와 호환되는 모든 애플리케이션과 통합하여 다양한 데이터 소스 및 데이터 형식에 걸쳐 통합된 메타데이터 리포지토리를 제공합니다. AWS Glue 크롤러는 Amazon S3의 소

스 데이터에서 스키마를 자동으로 추론하고 관련 메타데이터를 데이터 카탈로그에 저장할 수 있습니다. 데이터 카탈로그에 대한 자세한 내용은 AWS Glue 개발자 안내서의 Glue 데이터 카탈로그 채우기를 참조하세요. AWS

AWS Glue에는 별도의 요금이 적용됩니다. 데이터 카탈로그의 메타데이터 저장 및 액세스에 대한 월별 요금, Glue ETL 작업 및 AWS 크롤러 런타임에 대해 분당 청구되는 시간당 요금, 프로비저닝된 각 개발 엔드포인트에 대해 분당 청구되는 시간당 요금이 있습니다. 데이터 카탈로그에는 최대 100만 개의 객체를 무료로 저장할 수 있습니다. 100만 개 이상의 객체를 저장하는 경우, 100만 이상의 객체 100,000개마다 1USD가 부과됩니다. 데이터 카탈로그의 객체로는 테이블, 파티션 또는 데이터베이스가 있습니다. 자세한 내용은 Glue 요금을 참조하십시오.

> **Important**
>
> 2017년 8월 14일 이전에 Amazon Athena 또는 Amazon Redshift Spectrum을 사용하여 테이블을 생성한 경우 데이터베이스 및 테이블은 AWS Glue 데이터 카탈로그와 별도의 Athena 관리형 카탈로그에 저장됩니다. Amazon EMR을 이러한 테이블과 통합하려면 AWS Glue 데이터 카탈로그로 업그레이드해야 합니다. 자세한 내용은 Amazon Athena 사용 설명서의 AWS Glue 데이터 카탈로그로 업그레이드를 참조하세요.

Glue 데이터 카탈로그를 AWS 메타스토어로 지정

AWS Management Console AWS CLI또는 Amazon EMR API를 사용하여 AWS Glue 데이터 카탈로그를 메타스토어로 지정할 수 있습니다. CLI 또는 API를 사용할 때 Hive에 대한 구성 분류를 사용하여 데이터 카탈로그를 지정합니다. 또한 Amazon EMR 5.16.0 이상에서는 구성 분류를 사용하여 다른에서 데이터 카탈로그를 지정할 수 있습니다 AWS 계정. 콘솔을 사용할 때 고급 옵션 또는 빠른 옵션을 사용하여 데이터 카탈로그를 지정할 수 있습니다.

Console

콘솔을 사용하여 AWS Glue 데이터 카탈로그를 Hive 메타스토어로 지정하려면

1. 에 로그인 AWS Management Console하고 https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 왼쪽 탐색 창의 EMR on EC2에서 클러스터를 선택하고 클러스터 생성을 선택합니다.
3. 애플리케이션 번들에서 코어 Hadoop, HBase 또는 사용자 지정을 선택합니다. 클러스터를 사용자 지정하는 경우 Hive 또는 HCatalog를 애플리케이션 중 하나로 선택해야 합니다.
4. AWS Glue 데이터 카탈로그 설정에서 Hive 테이블 메타데이터에 사용 확인란을 선택합니다.

5. 클러스터에 적용할 다른 옵션을 선택합니다.
6. 클러스터를 시작하려면 클러스터 생성을 선택합니다.

CLI

를 사용하여 AWS Glue 데이터 카탈로그를 Hive 메타스토어로 지정하려면 AWS CLI

AWS CLI 및 EMR API를 사용하여 구성 분류를 지정하는 방법에 대한 자세한 내용은 섹션을 참조하세요[애플리케이션 구성](#).

- 다음 예에 표시된 대로 `hive-site` 구성 분류를 사용하여 `hive.metastore.client.factory.class`에 대한 값을 지정합니다.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

EMR 릴리스 버전 5.28.0, 5.28.1, 5.29.0 또는 6.x에서 AWS Glue 데이터 카탈로그를 메타스토어로 사용하여 클러스터를 생성하는 경우를 `hive.metastore.schema.verification`로 설정합니다`false`. 이렇게 하면 Hive 및 HCatalog가 MySQL에 대한 메타스토어 스키마를 검증하는 것을 방지합니다. 이러한 구성이 없으면, Hive 또는 HCatalog에서 재구성한 후에 프라이머리 인스턴스 그룹이 일시 중지됩니다.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false"
    }
  }
]
```

```
```

EMR 릴리스 버전 5.28.0, 5.28.1 또는 5.29.0에 이미 클러스터가 있는 경우, 다음 정보를 사용하여 프라이머리 인스턴스 그룹 `hive.metastore.schema.verification`을 `false`로 설정할 수 있습니다.

```
Classification = hive-site
Property       = hive.metastore.schema.verification
Value          = false
```

다른 AWS 계정에서 데이터 카탈로그를 지정하려면 다음 예제와 같이 `hive.metastore.glue.catalogid` 속성을 추가합니다. *`acct-id`*를 데이터 카탈로그의 AWS 계정으로 바꿉니다.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.schema.verification": "false",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

IAM 권한

클러스터의 EC2 인스턴스 프로파일에는 AWS Glue 작업에 대한 IAM 권한이 있어야 합니다. 또한 AWS Glue Data Catalog 객체에 대한 암호화를 활성화하는 경우 역할은 암호화에 AWS KMS key 사용되는를 암호화, 복호화 및 생성할 수 있어야 합니다.

AWS Glue 작업에 대한 권한

Amazon EMR에 대한 기본 EC2 인스턴스 프로파일을 사용할 경우 어떤 작업도 필요하지 않습니다. 에 연결된 `AmazonElasticMapReduceforEC2Role` 관리형 정책은 필요한 모든 AWS Glue 작업을 `EMR_EC2_DefaultRole` 허용합니다. 그러나 사용자 지정 EC2 인스턴스 프로파일 및 권한을 지정하하

는 경우 적절한 AWS Glue 작업을 구성해야 합니다. `AmazonElasticMapReduceforEC2Role` 관리형 정책을 시작점으로 사용합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 EC2 인스턴스의 서비스 역할(EC2 인스턴스 프로파일)을 참조하세요.

AWS Glue 데이터 카탈로그 암호화 및 복호화 권한

인스턴스 프로파일에 키를 사용하여 데이터를 암호화하고 복호화할 수 있는 권한이 필요합니다. 다음 두 설명이 모두 적용되는 경우에는 이러한 권한을 구성할 필요가 없습니다.

- AWS Glue용 관리형 키를 사용하여 AWS Glue Data Catalog 객체에 대한 암호화를 활성화합니다.
- Glue 데이터 카탈로그 AWS 계정 와 동일한 AWS 에 있는 클러스터를 사용합니다.

그렇지 않으면 EC2 인스턴스 프로파일에 연결된 권한 정책에 다음 명령을 추가해야 합니다.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Glue 데이터 카탈로그 암호화에 대한 자세한 내용은 AWS AWS Glue 개발자 안내서의 데이터 카탈로그 암호화를 참조하세요.

리소스 기반 권한

Amazon EMR에서 AWS Glue를 Hive, Spark 또는 Presto와 함께 사용하는 경우 AWS Glue는 데이터 카탈로그 리소스에 대한 액세스를 제어하는 리소스 기반 정책을 지원합니다. 이러한 리소스에는 데이터베이스, 테이블, 연결 및 사용자 정의 기능이 포함됩니다. 자세한 내용은 AWS Glue 개발자 안내서에서 AWS Glue 리소스 정책을 참조하세요.

리소스 기반 정책을 사용하여 Amazon EMR 내에서 AWS Glue에 대한 액세스를 제한하는 경우 권한 정책에서 지정하는 보안 주체는 클러스터 생성 시 지정된 EC2 인스턴스 프로파일과 연결된 역할 ARN이어야 합니다. 예를 들어, 카탈로그에 연결된 리소스 기반 정책의 경우 다음 예에 표시된 형식을 사용하여 클러스터 EC2 인스턴스에 대한 기본 서비스 역할의 ARN(*EMR_EC2_DefaultRole*)을 `Principal`로 지정할 수 있습니다.

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*는 AWS Glue 계정 ID와 다를 수 있습니다. 그러면 다른 계정의 EMR 클러스터에서 액세스할 수 있습니다. 각각 다른 계정에서 여러 보안 주체를 지정할 수 있습니다.

AWS Glue 데이터 카탈로그 사용 시 고려 사항

Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용할 때는 다음 항목을 고려하세요.

- Hive 셸을 사용한 보조 JAR 추가는 지원되지 않습니다. 차선책으로, `hive-site` 구성 분류를 사용하여 `hive.aux.jars.path` 속성을 설정하면 Hive 클래스 경로에 보조 JAR이 추가됩니다.
- Hive 트랜잭션은 지원되지 않습니다.
- AWS Glue 내에서 테이블 이름을 바꾸는 것은 지원되지 않습니다.
- `LOCATION`을 지정하지 않고 Hive 테이블을 생성하면 테이블 데이터가 `hive.metastore.warehouse.dir` 속성을 통해 지정된 위치에 저장됩니다. 기본적으로 HDFS에 있는 위치입니다. 다른 클러스터가 테이블에 액세스해야 하는 경우, 테이블을 생성한 클러스터에 대한 적절한 권한이 없으면 실패합니다. 또한 HDFS 스토리지는 일시적이기 때문에, 클러스터가 종료되는 경우 테이블 데이터가 손실되고 테이블을 다시 생성해야 합니다. Glue를 사용하여 Hive 테이블을 생성할 때 Amazon S3`LOCATION`에서 AWS 를 지정하는 것이 좋습니다. 또는 `hive-site` 구성 분류를 사용하여 모든 Hive 테이블에 적용되는 Amazon S3에서 `hive.metastore.warehouse.dir`에 대한 위치를 지정할 수도 있습니다. HDFS 위치에 테이블이 생성되고 테이블을 생성한 클러스터가 여전히 실행 중인 경우 AWS Glue 내에서 테이블 위치를 Amazon S3로 업데이트할 수 있습니다. 자세한 내용은 Glue 개발자 안내서의 AWS Glue 콘솔에서 테이블 작업을 참조하세요. AWS
- 따옴표와 아포스트로피가 포함된 파티션 값은 지원되지 않습니다(예: `PARTITION (owner="Doe's").`).
- 열 통계는 emr-5.31.0 이상에서 지원됩니다.
- Hive 권한 부여 사용은 지원되지 않습니다. 대안으로 AWS Glue 리소스 기반 정책 사용을 고려합니다. 자세한 내용은 AWS Glue 데이터 카탈로그에 대한 Amazon EMR 액세스를 위한 리소스 기반 정책 사용을 참조하세요.

- Hive 제약은 지원되지 않습니다.
- Hive의 비용 기반 최적화는 지원되지 않습니다.
- 설정 `hive.metastore.partition.inherit.table.properties`는 지원되지 않습니다.
- 다음 메타스토어 상수 사용은 지원되지 않습니다. `BUCKET_COUNT, BUCKET_FIELD_NAME, DDL_TIME, FIELD_TO_DIMENSION, FILE_INPUT_FORMAT, FILE_OUTPUT_FORMAT, HIVE_FILTER_FIELD_LAST_ACCESS, HIVE_FILTER_FIELD_OWNER, HIVE_FILTER_FIELD_PARAMS, IS_ARCHIVED, META_TABLE_COLUMNS, META_TABLE_COLUMN_TYPES, META_TABLE_DB, META_TABLE_LOCATION, META_TABLE_NAME, META_TABLE_PARTITION_COLUMNS, META_TABLE_SERDE, META_TABLE_STORAGE, ORIGINAL_LOCATION`.
- 조건자 표현식을 사용하는 경우에는 명시적 값이 비교 연산자 오른쪽에 와야 하며 그렇지 않으면 쿼리가 실패할 수 있습니다.
  - 올바름: `SELECT * FROM mytable WHERE time > 11`
  - 잘못됨: `SELECT * FROM mytable WHERE 11 > time`
- Amazon EMR 버전 5.32.0 및 6.3.0 이상에서는 조건자 표현식에서 사용자 정의 함수(UDF) 사용을 지원합니다. 이전 버전을 사용하는 경우 Hive가 쿼리 실행을 최적화하려는 방식 때문에 쿼리가 실패할 수 있습니다.
- 임시 테이블은 지원되지 않습니다.
- AWS Glue를 사용하여 직접 생성하는 대신 Amazon EMR을 통해 애플리케이션을 사용하여 테이블을 생성하는 것이 좋습니다. AWS Glue를 통해 테이블을 생성하면 필수 필드가 누락되고 쿼리 예외가 발생할 수 있습니다.
- EMR 5.20.0 이상에서는 AWS Glue 데이터 카탈로그를 메타스토어로 사용할 때 Spark 및 Hive에 대해 병렬 파티션 정리가 자동으로 활성화됩니다. 이 변경 사항은 여러 요청을 병렬로 실행하여 파티션을 검색함으로써 쿼리 계획 시간을 크게 단축합니다. 동시에 실행할 수 있는 총 세그먼트 수는 1~10입니다. 기본값은 5이며, 또한 권장 설정입니다. `hive-site` 구성 분류에서 `aws.glue.partition.num.segments` 속성을 지정하여 이 설정을 변경할 수 있습니다. 조절이 발생할 경우 값을 1로 변경하여 기능을 끌 수 있습니다. 자세한 내용은 AWS Glue 그먼트 구조를 참조하세요.

외부 MySQL 데이터베이스 또는 Amazon Aurora 사용

외부 MySQL 데이터베이스 또는 Amazon Aurora를 Hive 메타스토어로 사용하려면 Amazon RDS MySQL 인스턴스 또는 Amazon Aurora PostgreSQLinstance 인스턴스에 외부 데이터베이스 위치를 지정하도록 Hive의 메타스토어에 대한 기본 구성 값을 재정의합니다.

> **Note**
>
> Hive는 메타스토어 테이블에 대한 동시 쓰기 액세스를 지원하지 않으며 금지하지도 않습니다. 두 클러스터 간에 메타스토어 정보를 공유하는 경우, 동일한 메타스토어 테이블의 다른 파티션에 쓰고 있지 않은 한, 동시에 동일한 메타스토어 테이블에 쓰지 않는지 확인해야 합니다.

다음 절차에서는 Hive 메타스토어 위치에 대한 기본 구성 값을 재정의하고 재구성된 메타스토어 위치를 사용하여 클러스터를 시작하는 방법을 보여 줍니다.

EMR 클러스터 외부에 위치한 메타스토어를 생성하려면

1. MySQL 또는 Aurora PostgreSQL 데이터베이스를 생성합니다. PostgreSQL을 사용하는 경우 클러스터를 프로비저닝한 후에 구성해야 합니다. 클러스터 생성 시 MySQL만 지원됩니다. Aurora MySQL 및 Aurora PostgreSQL 간 차이점에 대한 자세한 내용은 [Amazon Aurora MySQL의 개요](#) 및 [Amazon Aurora PostgreSQL 작업](#)을 참조하세요. 일반적으로 Amazon RDS 데이터베이스를 생성하는 방법에 대한 자세한 내용은 https://aws.amazon.com/rds/를 참조하세요.
2. 데이터베이스와 ElasticMapReduce-Master 보안 그룹 간의 JDBC 연결을 허용하도록 보안 그룹을 수정합니다. 액세스를 위한 보안 그룹을 수정하는 방법에 대한 자세한 내용은 [Amazon EMR 관리형 보안 그룹 작업](#)을 참조하세요.
3. `hive-site.xml`에서 JDBC 구성 값을 설정합니다.

   > **Important**
   >
   > 암호와 같은 중요한 정보를 Amazon EMR 구성 API에 제공하면 해당 정보가 충분한 권한이 있는 계정에 표시됩니다. 이 정보가 다른 사용자에게 표시될 수 있다고 우려되는 경우 관리 계정으로 클러스터를 생성하고 `elasticmapreduce:DescribeCluster` API 키에 대한 권한을 명시적으로 거부하는 역할을 생성하여 다른 사용자(IAM 사용자 또는 위임된 자격 증명이 있는 사용자)를 클러스터의 서비스 액세스로 제한합니다.

   a. 다음 예제에 표시된 대로 `hiveConfiguration.json`에 대한 편집 내용이 포함된 `hive-site.xml`이라는 구성 파일을 생성합니다.

      *`hostname`*을 데이터베이스를 실행하는 Amazon RDS 인스턴스의 DNS 주소로 바꾸고 *`username`* 및 *`password`*를 데이터베이스의 보안 인증으로 바꿉니다. MySQL 및 Aurora 데이터베이스 인스턴스에 연결하는 방법에 대한 자세한 내용은 Amazon RDS 사용 설명서에

서 [MySQL 데이터베이스 엔진을 실행하는 DB 인스턴스에 연결](#) 및 [Athena DB 클러스터에 연결](#)을 참조하세요. `javax.jdo.option.ConnectionURL`은 JDBC 메타스토어의 JDBC 연결 문자열이고, `javax.jdo.option.ConnectionDriverName`은 JDBC 메타스토어의 드라이버 클래스 이름입니다.

MySQL JDBC 드라이버는 Amazon EMR에 의해 설치됩니다.

값 속성에 공백이나 캐리지 리턴을 포함할 수 없습니다. 속성은 모두 한 줄에 나타나야 합니다.

```
[
    {
      "Classification": "hive-site",
      "Properties": {
        "javax.jdo.option.ConnectionURL": "jdbc:mysql://hostname:3306/hive?
createDatabaseIfNotExist=true",
        "javax.jdo.option.ConnectionDriverName": "org.mariadb.jdbc.Driver",
        "javax.jdo.option.ConnectionUserName": "username",
        "javax.jdo.option.ConnectionPassword": "password"
      }
    }
  ]
```

b. 다음 AWS CLI 명령과 같이 클러스터를 생성할 때 `hiveConfiguration.json` 파일을 참조합니다. 이 명령에서 파일은 로컬로 저장되며, 파일을 Amazon S3에 업로드하고 여기에서 참조할 수도 있습니다(예: `s3://DOC-EXAMPLE-BUCKET/hiveConfiguration.json`).

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --release-label emr-7.9.0 --instance-type m5.xlarge --
instance-count 2 \
--applications Name=Hive --configurations file://hiveConfiguration.json --use-
default-roles
```

4. 클러스터의 프라이머리 노드에 연결합니다.

SSH를 사용하여 프라이머리 노드에 연결하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.

5. 다음과 같은 명령을 입력하여 Amazon S3의 위치를 지정하는 Hive 테이블을 생성합니다.

```
CREATE EXTERNAL TABLE IF NOT EXISTS table_name
(
key int,
value int
)
LOCATION s3://DOC-EXAMPLE-BUCKET/hdfs/
```

6. 실행 중인 클러스터에 Hive 스크립트를 추가합니다.

Hive 클러스터는 Amazon RDS에 위치한 메타스토어를 사용하여 실행됩니다. 메타스토어 위치를 지정하여 이 메타스토어를 공유하는 모든 추가 Hive 클러스터를 시작합니다.

Hive JDBC 드라이버 사용

Microsoft Excel, MicroStrategy, QlikView, Tableau 등과 같은 인기 있는 비즈니스 인텔리전스 도구를 Amazon EMR과 함께 사용하여 데이터를 탐색하고 시각화할 수 있습니다. 이러한 도구에는 대부분 ODBC(Open Database Connectivity) 또는 JDBC(Java Database Connectivity) 드라이버가 필요합니다. Amazon EMR은 JDBC 및 ODBC 연결을 모두 지원합니다.

아래 예제는 SQL Workbench/J를 SQL 클라이언트로 사용하여 Amazon EMR의 Hive 클러스터에 연결하는 방법을 보여줍니다. 기타 드라이버의 경우 Amazon EMR에서 비즈니스 인텔리전스 도구 사용을 참조하세요.

SQL Workbench/J를 설치하여 사용하기 전에 드라이버 패키지를 다운로드하여 드라이버를 설치하십시오. 패키지에 포함된 드라이버는 Amazon EMR 릴리스 버전 4.0 이상에서 사용 가능한 Hive 버전을 지원합니다. 자세한 출시 정보 및 설명서를 보려면 패키지에 포함된 PDF 설명서를 참조하십시오.

- 최신 Hive JDBC 드라이버 패키지 다운로드

  http://awssupportdatasvcs.com/bootstrap-actions/Simba/latest/
- 이전 버전의 Hive JDBC 드라이버

  http://awssupportdatasvcs.com/bootstrap-actions/Simba/

SQL Workbench를 설치하고 구성하려면

1. http://www.sql-workbench.net/downloads.html에서 운영 체제에 맞는 SQL Workbench/J 클라이언트를 다운로드합니다.
2. 자세한 내용은 SQL Workbench/J 사용 설명서에서 Installing and starting SQL Workbench/J를 참조하십시오.
3. Linux, Unix, Mac OS X 사용자: 터미널 세션에서 다음 명령을 사용하여 클러스터의 마스터 노드에 대한 SSH 터널을 생성합니다. *master-public-dns-name*을 프라이머리 노드의 퍼블릭 DNS 이름으로 바꾸고 *path-to-key-file*을 Amazon EC2 프라이빗 키(`.pem`) 파일의 위치 및 파일 이름으로 바꿉니다.

```
ssh -o ServerAliveInterval=10 -i path-to-key-file -N -L 10000:localhost:10000
 hadoop@master-public-dns-name
```

   Windows 사용자: PuTTY 세션에서 소스 포트에 `10000`, 대상에 *master-public-dns-name*`:10000`을 사용하여 클러스터의 프라이머리 노드에 대한 SSH 터널을 생성합니다(로컬 포트 전달 사용). *master-public-dns-name*을 마스터 노드의 퍼블릭 DNS 이름으로 바꿉니다.
4. JDBC 드라이버를 SQL Workbench에 추가합니다.

   a. Select Connection Profile(연결 프로파일 선택) 대화 상자에서 Manage Drivers(드라이버 관리)를 클릭합니다.
   b. Create a new entry(새 항목 생성)(빈 페이지) 아이콘을 클릭합니다.
   c. Name(이름) 필드에 **`Hive JDBC`**을(를) 입력합니다.
   d. 라이브러리에서 Select the JAR file(s)(JAR 파일 선택) 아이콘을 클릭합니다.
   e. 추출된 드라이버를 포함하는 위치로 이동합니다. 다운로드한 JDBC 드라이버 패키지 버전에 포함된 드라이버를 선택하고 열기를 클릭합니다.

      예를 들어 JDBC 드라이버 패키지에는 다음 JAR이 포함될 수 있습니다.

```
hive_metastore.jar
hive_service.jar
HiveJDBC41.jar
libfb303-0.9.0.jar
libthrift-0.9.0.jar
log4j-1.2.14.jar
ql.jar
slf4j-api-1.5.11.jar
slf4j-log4j12-1.5.11.jar
```

```
TCLIServiceClient.jar
zookeeper-3.4.6.jar
```

   f. Please select one driver(드라이버 하나를 선택하십시오) 대화 상자에서 `com.amazon.hive.jdbc41.HS2Driver`와 OK(확인)를 선택합니다.

5. Manage Drivers(드라이버 관리) 대화 상자로 돌아가면 Classname 필드가 채워져 있는지 확인하고 OK(확인)를 선택합니다.
6. Select Connection Profile(연결 프로파일 선택) 대화 상자로 돌아가면 Driver(드라이버) 필드가 Hive JDBC로 설정되어 있는지 확인하고 URL 필드에 `jdbc:hive2://localhost:10000/default` JDBC 연결 문자열을 제공합니다.
7. OK(확인)를 선택하여 연결합니다. 연결이 완료된 후 SQL Workbench/J 창 상단에 연결 세부 정보가 나타납니다.

Hive 및 JDBC 인터페이스 사용에 대한 자세한 내용은 Apache Hive 설명서의 HiveClient 및 HiveJDBCInterface를 참조하십시오.

Hive 성능 개선

Amazon EMR은 Amazon S3에 저장된 데이터를 쿼리, 읽기 및 쓰기 위해 Hive를 사용할 때 성능을 최적화하는 기능을 제공합니다.

S3 Select는 Amazon S3로 처리를 '푸시다운'하여 일부 애플리케이션의 CSV 및 JSON 파일에 대한 쿼리 성능을 향상시킬 수 있습니다.

EMRFS S3 최적화 커미터는 OutputCommitter 클래스의 대안으로, EMRFS를 사용하여 Amazon S3에서 파일을 작성할 때 나열 및 이름 바꾸기 작업을 제거하여 성능을 개선합니다.

주제

Hive EMRFS S3 최적화 커미터 활성화

Hive EMRFS S3 최적화 커미터는 EMRFS를 사용할 때 EMR Hive가 삽입 쿼리용 파일을 작성하는 데 사용하는 대체 방법입니다. 커미터는 Amazon S3에서 수행된 나열 및 이름 바꾸기 작업을 없애고 애플리케이션 성능을 개선합니다. 이 기능은 EMR 5.34 및 EMR 6.5부터 사용할 수 있습니다.

커미터 활성화

EMR Hive가 `HiveEMRFSOptimizedCommitter`를 사용하여 모든 Hive 관리형 테이블 및 외부 테이블의 기본값으로 데이터를 커밋할 수 있도록 하려면 EMR 6.5.0 또는 EMR 5.34.0 클러스터에서 다음 `hive-site` 구성을 사용합니다.

```
[
   {
      "classification": "hive-site",
      "properties": {
         "hive.blobstore.use.output-committer": "true"
      }
   }
]
```

Note

`hive.exec.parallel`이 `true`로 설정된 경우 이 기능을 켜지 않습니다.

제한 사항

태그에 적용되는 기본 제한은 다음과 같습니다.

- Hive에서 작은 파일의 자동 병합 기능은 지원되지 않습니다. 최적화된 커미터가 활성화된 경우에도 기본 Hive 커밋 로직이 사용됩니다.
- Hive ACID 테이블은 지원되지 않습니다. 최적화된 커미터가 활성화된 경우에도 기본 Hive 커밋 로직이 사용됩니다.
- Hive에서 작성된 파일의 파일 이름 지정 방식이 `<task_id>_<attempt_id>_<copy_n>`에서 `<task_id>_<attempt_id>_<copy_n>_<query_id>`로 변경되었습니다. 예를 들어

  `s3://warehouse/table/partition=1/000000_0` 파일은 `s3://warehouse/table/partition=1/000000_0-hadoop_20210714130459_ba7c23ec-5695-4947-9d98-8a40ef759222-1`로 변경됩니다. 여기서 `query_id`는 사용자 이름, 타임스탬프 및 UUID의 조합입니다.
- 사용자 지정 파티션이 다른 파일 시스템(HDFS, S3)에 있는 경우 이 기능은 자동으로 비활성화됩니다. 활성화되면 기본 Hive 커밋 로직이 사용됩니다.

S3 Select와 함께 Hive를 사용하여 성능 향상

> **Important**
>
> 신규 고객은 더 이상 Amazon S3 Select를 사용할 수 없습니다. Amazon S3 Select 기존 고객은 평소처럼 이 기능을 계속 사용할 수 있습니다. 자세히 알아보기

Amazon EMR 릴리스 버전 5.18.0 이상에서는 Amazon EMR 기반 Hive에서 S3 Select를 사용할 수 있습니다. S3 Select를 사용하면 애플리케이션이 객체에서 데이터 하위 집합만 검색할 수 있습니다. Amazon EMR의 경우 처리를 위해 대형 데이터 세트를 필터링하는 계산 작업이 클러스터에서 Amazon S3로 '푸시다운'되어 일부 애플리케이션의 성능이 향상되고 Amazon EMR과 Amazon S3 사이에 전송되는 데이터의 양이 줄어듭니다.

S3 Select에는 CSV 및 JSON 파일을 기반으로, 그리고 Hive 세션 중에 `s3select.filter` 구성 변수를 `true`로 설정하여 Hive 테이블이 지원됩니다. 자세한 정보와 지침은 코드에서 S3 Select 지정 섹션을 참조하세요.

S3 Select가 애플리케이션에 적합한가요?

S3 Select 사용 여부에 관계없이 애플리케이션을 벤치마킹하여 애플리케이션에 적합한지 여부를 확인하는 것이 좋습니다.

다음 지침을 사용하여 애플리케이션과 함께 S3 Select를 사용할 수 있는지 확인하십시오.

- 쿼리가 원본 데이터 세트 중 반 이상을 필터링합니다.
- 쿼리 필터 조건자가 Amazon S3 Select에서 지원하는 데이터 형식의 열을 사용합니다. 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 데이터 유형을 참조하세요.
- Amazon S3 및 Amazon EMR 클러스터 간 네트워크 연결의 전송 속도와 가용 대역폭이 양호합니다. Amazon S3는 HTTP 응답을 압축하지 않으므로 압축된 입력 파일의 경우 응답 크기가 증가할 수 있습니다.

고려 사항 및 제한 사항

- 고객 제공 암호화 키를 사용하는 Amazon S3 서버 측 암호화(SSE-C) 및 클라이언트 측 암호화는 지원되지 않습니다.
- `AllowQuotedRecordDelimiters` 속성이 지원되지 않습니다. 이 속성이 지정되면 쿼리가 실패합니다.

- UTF-8 형식의 CSV 및 JSON 파일만 지원됩니다. 여러 줄의 CSV 및 JSON은 지원되지 않습니다.
- 압축되지 않거나 gzip 또는 bzip2 파일이 지원되지 않습니다.
- 마지막 줄의 설명 문자는 지원되지 않습니다.
- 파일 끝의 빈 줄은 처리되지 않습니다.
- Amazon EMR 기반 Hive는 S3 Select에서 지원하는 기본 데이터 형식을 지원합니다. 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 데이터 유형을 참조하세요.

코드에서 S3 Select 지정

Hive 테이블에 S3 Select를 사용하려면 `com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat`을 INPUTFORMAT 클래스 이름으로 지정하여 테이블을 생성하고, TBLPROPERTIES 절을 사용하여 `s3select.format` 속성의 값을 지정합니다.

쿼리를 실행할 때 기본적으로 S3 Select가 비활성화됩니다. 아래와 같이 Hive에서 `s3select.filter`를 `true`로 설정하여 S3 Select를 활성화합니다. 아래의 예제는 기본 CSV 및 JSON 파일로 테이블을 생성한 후 간단한 select 문을 사용하여 테이블을 쿼리할 때 S3 Select를 지정하는 방법을 보여 줍니다.

Example CSV 기반 테이블의 CREATE TABLE 문

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
STORED AS
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/mycsvfile/'
TBLPROPERTIES (
  "s3select.format" = "csv",
  "s3select.headerInfo" = "ignore"
);
```

Example JSON 기반 테이블의 CREATE TABLE 문

```
CREATE TABLE mys3selecttable (
col1 string,
col2 int,
col3 boolean
)
ROW FORMAT SERDE 'org.apache.hive.hcatalog.data.JsonSerDe'
STORED AS
INPUTFORMAT
  'com.amazonaws.emr.s3select.hive.S3SelectableTextInputFormat'
OUTPUTFORMAT
  'org.apache.hadoop.hive.ql.io.HiveIgnoreKeyTextOutputFormat'
LOCATION 's3://path/to/json/'
TBLPROPERTIES (
  "s3select.format" = "json"
);
```

Example SELECT TABLE 문

```
SET s3select.filter=true;
SELECT * FROM mys3selecttable WHERE col2 > 10;
```

MSCK 최적화

Hive는 해당 메타스토어에 각 테이블의 파티션 목록을 저장합니다. 그러나 파티션이 파일 시스템에 직접 추가되거나 파일 시스템에서 제거되는 경우 Hive 메타스토어는 이러한 변경 사항을 인식하지 못합니다. MSCK 명령은 파일 시스템에서 직접 추가 또는 제거된 파티션에 대해 Hive 메타스토어에 있는 파티션 메타데이터를 업데이트합니다. 명령 구문은 다음과 같습니다.

```
MSCK [REPAIR] TABLE table_name [ADD/DROP/SYNC PARTITIONS];
```

Hive는 이 명령을 다음과 같이 구현합니다.

1. Hive는 메타스토어에서 테이블의 모든 파티션을 검색합니다. 파일 시스템에 없는 파티션 경로 목록에서 메타스토어로부터 삭제할 파티션 목록을 생성합니다.
2. Hive는 파일 시스템에 있는 파티션 경로를 수집하여 메타스토어의 파티션 목록과 비교하고 메타스토어에 추가해야 하는 파티션 목록을 생성합니다.
3. Hive는 ADD, DROP 또는 SYNC 모드를 사용하여 메타스토어를 업데이트합니다.

> **Note**
>
> 메타스토어에 파티션이 많은 경우 각 파티션에 대해 파일 시스템의 `exists` API 직접 호출을 수행해야 하므로 파일 시스템에 파티션이 없는지 확인하는 단계를 실행하는 데 시간이 오래 걸립니다.

Amazon EMR 6.5.0에서 Hive는 `hive.emr.optimize.msck.fs.check`라는 플래그를 도입했습니다. 이 플래그가 활성화되면 Hive는 파일 시스템 API를 직접 호출하는 대신 위의 2단계에서 생성된 파일 시스템의 파티션 경로 목록에서 파티션이 있는지 확인합니다. Amazon EMR 6.8.0에서 Hive는 기본적으로 이 최적화 기능을 활성화하므로 `hive.emr.optimize.msck.fs.check` 플래그를 설정할 필요가 없습니다.

Hive Live Long and Process(LLAP) 사용

Amazon EMR 6.0.0은 Hive용 LLAP(Live Long and Process) 기능을 지원합니다. LLAP는 지능형 인 메모리 캐싱을 통해 영구 데몬을 사용하여 이전의 기본 Tez 컨테이너 실행 모드에 비해 쿼리 성능을 향상시킵니다.

Hive LLAP 데몬은 YARN 서비스로 관리되고 실행됩니다. YARN 서비스는 장기 실행 YARN 애플리케이션으로 간주될 수 있으므로 일부 클러스터 리소스가 Hive LLAP 전용으로 할당되고 다른 워크로드에 사용될 수 없습니다. 자세한 내용은 LLAP 및 YARN 서비스 API를 참조하십시오.

Amazon EMR에서 Hive LLAP 활성화

Amazon EMR에서 Hive LLAP를 활성화하려면 클러스터를 시작할 때 다음 구성을 제공합니다.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true"
    }
  }
]
```

자세한 내용은 애플리케이션 구성을 참조하세요.

기본적으로 Amazon EMR은 클러스터 YARN 리소스의 약 60%를 Hive LLAP 대몬(daemon)에 할당합니다. Hive LLAP에 할당되는 클러스터 YARN 리소스의 비율과 Hive LLAP 할당에 고려할 작업 및 코어 노드의 수를 구성할 수 있습니다.

예를 들어, 다음 구성은 3개의 작업 또는 코어 노드에서 3개의 데몬으로 Hive LLAP를 시작하고 3개의 코어 또는 작업 노드의 YARN 리소스 중 40%를 Hive LLAP 데몬에 할당합니다.

```
[
  {
    "Classification": "hive",
    "Properties": {
      "hive.llap.enabled": "true",
      "hive.llap.percent-allocation": "0.4",
      "hive.llap.num-instances": "3"
    }
  }
]
```

분류 API에서 다음 `hive-site` 구성을 사용하여 기본 LLAP 리소스 설정을 재정의할 수 있습니다.

속성	설명
hive.llap.daemon.yarn.container.mb	총 LLAP 데몬 컨테이너 크기(MB)
hive.llap.daemon.memory.per.instance.mb	LLAP 데몬 컨테이너의 실행 프로그램이 사용하는 총 메모리(MB)
hive.llap.io.memory.size	LLAP 입력/출력의 캐시 크기
hive.llap.daemon.num.executors	LLAP 데몬당 실행 프로그램 수

클러스터에서 Hive LLAP 수동으로 시작

LLAP에서 사용하는 모든 종속 항목 및 구성은 클러스터 시작 과정에서 LLAP tar 아카이브로 패키징됩니다. `"hive.llap.enabled": "true"`를 사용하여 LLAP를 활성화한 경우 Amazon EMR 재구성을 사용하여 LLAP의 구성을 변경하는 것이 좋습니다.

이렇게 하지 않고 `hive-site.xml`을 수동으로 변경하려면, 다음 예제와 같이 `hive --service llap` 명령을 사용하여 LLAP tar 아카이브를 다시 빌드해야 합니다.

```
# Define how many resources you want to allocate to Hive LLAP

LLAP_INSTANCES=<how many llap daemons to run on cluster>
LLAP_SIZE=<total container size per llap daemon>
LLAP_EXECUTORS=<number of executors per daemon>
LLAP_XMX=<Memory used by executors>
LLAP_CACHE=<Max cache size for IO allocator>

yarn app -enableFastLaunch

hive --service llap \
--instances $LLAP_INSTANCES \
--size ${LLAP_SIZE}m \
--executors $LLAP_EXECUTORS \
--xmx ${LLAP_XMX}m \
--cache ${LLAP_CACHE}m \
--name llap0 \
--auxhbase=false \
--startImmediately
```

Hive LLAP 상태 확인

Hive를 통해 Hive LLAP의 상태를 확인하려면 다음 명령을 사용합니다.

```
hive --service llapstatus
```

YARN을 사용하여 Hive LLAP의 상태를 확인하려면 다음 명령을 사용합니다.

```
yarn app -status (name-of-llap-service)

# example:
yarn app -status llap0 | jq
```

Hive LLAP 시작 또는 중지

Hive LLAP는 영구 YARN 서비스로 실행되므로 YARN 서비스를 중지하거나 다시 시작하여 Hive LLAP를 중지하거나 다시 시작합니다. 다음 명령은 이를 보여줍니다.

```
yarn app -stop llap0
yarn app -start llap0
```

Hive LLAP 대몬(daemon) 수 조정

LLAP 인스턴스 수를 줄이려면 다음 명령을 사용합니다.

```
yarn app -flex llap0 -component llap -1
```

자세한 내용은 서비스 구성 요소 Flex를 참조하십시오.

Hive에서 암호화

이 섹션에서는 Amazon EMR에서 지원하는 암호화 유형을 설명합니다.

Hive에서 Parquet 모듈식 암호화

Parquet 모듈식 암호화는 열 수준 액세스 제어 및 암호화를 제공하여 Parquet 파일 형식으로 저장된 데이터의 개인 정보 보호 및 데이터 무결성을 개선합니다. 이 기능은 릴리스 6.6.0부터 Amazon EMR Hive에서 사용할 수 있습니다.

파일 암호화 또는 스토리지 계층 암호화를 포함하여 이전에 지원되는 보안 및 무결성 솔루션은 Amazon EMR 관리 안내서의 암호화 옵션에 설명합니다. 이러한 솔루션을 Parquet 파일에 사용할 수 있지만, 통합된 Parquet 암호화 메커니즘의 새로운 기능을 활용하면 열 수준에 대한 세분화된 액세스가 가능하고 성능 및 보안이 개선됩니다. Apache github 페이지 Parquet Modular Encryption에서 이 기능에 대해 자세히 알아보세요.

사용자는 Hadoop 구성을 사용하여 Parquet 리더와 라이터에 구성을 전달합니다. 사용자가 암호화를 활성화하도록 리더와 라이터를 구성하고 고급 기능을 전환하는 방법에 대한 자세한 구성은 PARQUET-1854: Properties-driven Interface to Parquet Encryption Management에 나와 있습니다.

사용 예제:

다음 예제에서는 암호화 키를 관리하는 데 AWS KMS 를 사용하여 Hive 테이블을 생성하고 여기에 데이터를 작성하는 방법을 설명합니다.

1. PARQUET-1373: Encryption Key Management Tools 문서에 설명된 대로 AWS KMS 서비스에 대한 KmsClient를 구현합니다. 다음 샘플은 구현 스니펫을 보여줍니다.

```
package org.apache.parquet.crypto.keytools;
```

```
import com.amazonaws.AmazonClientException;
import com.amazonaws.AmazonServiceException;
import com.amazonaws.regions.Regions;
import com.amazonaws.services.kms.AWSKMS;
import com.amazonaws.services.kms.AWSKMSClientBuilder;
import com.amazonaws.services.kms.model.DecryptRequest;
import com.amazonaws.services.kms.model.EncryptRequest;
import com.amazonaws.util.Base64;
import org.apache.hadoop.conf.Configuration;
import org.apache.parquet.crypto.KeyAccessDeniedException;
import org.apache.parquet.crypto.ParquetCryptoRuntimeException;
import org.apache.parquet.crypto.keytools.KmsClient;
import org.slf4j.Logger;
import org.slf4j.LoggerFactory;

import java.nio.ByteBuffer;
import java.nio.charset.Charset;
import java.nio.charset.StandardCharsets;

public class AwsKmsClient implements KmsClient {

    private static final AWSKMS AWSKMS_CLIENT = AWSKMSClientBuilder
            .standard()
            .withRegion(Regions.US_WEST_2)
            .build();
    public static final Logger LOG = LoggerFactory.getLogger(AwsKmsClient.class);

    private String kmsToken;
    private Configuration hadoopConfiguration;

    @Override
    public void initialize(Configuration configuration, String kmsInstanceID, String
 kmsInstanceURL, String accessToken) throws KeyAccessDeniedException {
        hadoopConfiguration = configuration;
        kmsToken = accessToken;

    }

    @Override
    public String wrapKey(byte[] keyBytes, String masterKeyIdentifier) throws
 KeyAccessDeniedException {
        String value = null;
        try {
```

```
            ByteBuffer plaintext = ByteBuffer.wrap(keyBytes);

            EncryptRequest req = new
 EncryptRequest().withKeyId(masterKeyIdentifier).withPlaintext(plaintext);
            ByteBuffer ciphertext = AWSKMS_CLIENT.encrypt(req).getCiphertextBlob();

            byte[] base64EncodedValue = Base64.encode(ciphertext.array());
            value = new String(base64EncodedValue, Charset.forName("UTF-8"));
        } catch (AmazonClientException ae) {
            throw new KeyAccessDeniedException(ae.getMessage());
        }
        return value;
    }

    @Override
    public byte[] unwrapKey(String wrappedKey, String masterKeyIdentifier) throws
 KeyAccessDeniedException {
        byte[] arr = null;
        try {
            ByteBuffer ciphertext  =
 ByteBuffer.wrap(Base64.decode(wrappedKey.getBytes(StandardCharsets.UTF_8)));
            DecryptRequest request = new
 DecryptRequest().withKeyId(masterKeyIdentifier).withCiphertextBlob(ciphertext);
            ByteBuffer decipheredtext =
 AWSKMS_CLIENT.decrypt(request).getPlaintext();
            arr = new byte[decipheredtext.remaining()];
            decipheredtext.get(arr);
        } catch (AmazonClientException ae) {
            throw new KeyAccessDeniedException(ae.getMessage());
        }
        return arr;
    }
}
```

2. AWS Key Management Service 개발자 안내서의 키 생성에 설명된 대로 바닥글에 대한 AWS KMS 암호화 키와 IAM 역할이 액세스할 수 있는 열을 생성합니다. 기본 IAM 역할은 EMR_ECS_default입니다.
3. Amazon EMR 클러스터의 Hive 애플리케이션에서 Apache Hive 리소스 설명서에 설명된 대로 ADD JAR 문을 사용하여 위의 클라이언트를 추가합니다. 다음은 명령문 예제입니다.

```
ADD JAR 's3://location-to-custom-jar';
```

다른 방법은 부트스트랩 작업을 사용하여 JAR을 Hive의 `auxlib`에 추가하는 것입니다. 다음은 부트스트랩 작업에 추가할 예제 줄입니다.

```
aws s3 cp 's3://location-to-custom-jar' /usr/lib/hive/auxlib
```

4. 다음과 같은 구성을 설정합니다.

```
set
 parquet.crypto.factory.class=org.apache.parquet.crypto.keytools.PropertiesDrivenCryptoFacto
set
 parquet.encryption.kms.client.class=org.apache.parquet.crypto.keytools.AwsKmsClient;
```

5. Parquet 형식으로 Hive 테이블을 생성하고 SERDEPROPERTIES에서 AWS KMS 키를 지정하고 여기에 일부 데이터를 삽입합니다.

```
CREATE TABLE my_table(name STRING, credit_card STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
WITH SERDEPROPERTIES (
  'parquet.encryption.column.key’=<aws-kms-key-id-for-column-1>: credit_card’,
  'parquet.encryption.footer.key’='<aws-kms-key-id-for-footer>’)
STORED AS parquet
LOCATION “s3://<bucket/<warehouse-location>/my_table”;

INSERT INTO my_table SELECT
java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',5) as
 name,
java_method ('org.apache.commons.lang.RandomStringUtils','randomAlphabetic',10) as
 credit_card
from (select 1) x lateral view posexplode(split(space(100),' ')) pe as i,x;

select * from my_table;
```

6. AWS KMS 키에 대한 액세스 권한이 없는 동일한 위치에 외부 테이블을 생성할 때(예: IAM 역할 액세스가 거부됨) 데이터를 읽을 수 없는지 확인합니다.

```
CREATE EXTERNAL TABLE ext_table (name STRING, credit_card STRING)
ROW FORMAT SERDE 'org.apache.hadoop.hive.ql.io.parquet.serde.ParquetHiveSerDe’
STORED AS parquet
LOCATION “s3://<bucket>/<warehouse-location>/my_table”;

SELECT * FROM ext_table;
```

7. 마지막 명령문에서 다음 예외가 발생합니다.

```
Failed with exception
 java.io.IOException:org.apache.parquet.crypto.KeyAccessDeniedException: Footer key:
 access denied
```

HiveServer2에서 전송 중 데이터 암호화

Amazon EMR 릴리스 6.9.0부터 HiveServer2(HS2)는 HiveServer2에서 전송 중 데이터 암호화 보안 구성의 일부로 TLS 및 SSL을 지원합니다. 이는 전송 중 암호화가 활성화된 상태로 Amazon EMR 클러스터에서 실행 중인 HS2에 연결하는 방법에 영향을 줍니다. HS2에 연결하려면 JDBC URL의 TRUSTSTORE_PATH 및 TRUSTSTORE_PASSWORD 파라미터 값을 수정해야 합니다. 다음 URL은 필수 파라미터가 포함된 HS2용 JDBC 연결의 예제입니다.

```
jdbc:hive2://HOST_NAME:10000/
default;ssl=true;sslTrustStore=TRUSTSTORE_PATH;trustStorePassword=TRUSTSTORE_PASSWORD
```

아래의 클러스터 내부 또는 클러스터 외부 HiveServer2 암호화에 대한 적절한 지침을 사용합니다.

On-cluster HS2 access

프라이머리 노드에 SSH로 연결한 후 Beeline 클라이언트를 사용하여 HiveServer2에 액세스하는 경우 `ssl.server.truststore.location` 및 `ssl.server.truststore.password` 구성을 사용하여 TRUSTSTORE_PATH 및 TRUSTSTORE_PASSWORD 파라미터를 찾기 위해 `/etc/hadoop/conf/ssl-server.xml`을 참조합니다.

다음 예제 명령은 이러한 구성을 검색하는 데 도움이 될 수 있습니다.

```
TRUSTSTORE_PATH=$(sed -n '/ssl.server.truststore.location/,+2p' /etc/hadoop/conf/
ssl-server.xml | awk -F "[><]" '/value/{print $3}')
TRUSTSTORE_PASSWORD=$(sed -n '/ssl.server.truststore.password/,+2p' /etc/hadoop/
conf/ssl-server.xml | awk -F "[><]" '/value/{print $3}')
```

Off-cluster HS2 access

Amazon EMR 클러스터 외부의 클라이언트에서 HiveServer2에 액세스하는 경우 다음 접근 방식 중 하나를 사용하여 TRUSTSTORE_PATH 및 TRUSTSTORE_PASSWORD를 가져올 수 있습니다.

- 보안 구성 중에 생성된 PEM 파일을 JKS 파일로 변환하고 JDBC 연결 URL에서 동일한 파일을 사용합니다. 예를 들어 openssl 및 keytool을 사용하는 경우 다음 명령을 사용합니다.

```
openssl pkcs12 -export -in trustedCertificates.pem -inkey privateKey.pem -out
 trustedCertificates.p12 -name "certificate"
keytool -importkeystore -srckeystore trustedCertificates.p12 -srcstoretype pkcs12
 -destkeystore trustedCertificates.jks
```

- 또는 `ssl.server.truststore.location` 및 `ssl.server.truststore.password` 구성을 사용해 `/etc/hadoop/conf/ssl-server.xml`을 참조하여 `TRUSTSTORE_PATH` 및 `TRUSTSTORE_PASSWORD` 파라미터 값을 찾습니다. 트러스트 스토어 파일을 클라이언트 시스템에 다운로드하고 클라이언트 컴퓨터의 경로를 `TRUSTSTORE_PATH`로 사용합니다.

  Amazon EMR 클러스터 외부의 클라이언트에서 애플리케이션에 액세스하는 방법에 대한 자세한 내용은 Hive JDBC 드라이버 사용을 참조하세요.

Hive 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Hive의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Hive 버전 정보

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-7.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
		worker, zookeeper-client, zookeeper-server
emr-7.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-ya rn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-7.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.6.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-7.5.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.4.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-7.3.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-7.2.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.36.2	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-7.1.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-7.0.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.15.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.14.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.13.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.12.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.11.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.11.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.10.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.10.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, tez-on-worker, zookeeper-client, zookeeper-server
emr-6.9.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.9.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.8.1	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.8.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.7.0	3.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.36.1	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.36.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.6.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-5.35.0	2.3.9	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.5.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.4.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.3.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.3.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.2.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.2.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.1.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.1.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-6.0.1	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server
emr-6.0.0	3.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.34.0	2.3.8	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.33.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.33.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.32.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.32.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.31.1	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.31.0	2.3.7	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.30.2	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.30.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn
emr-5.30.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mariadb-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.29.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn
emr-5.28.1	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.28.0	2.3.6	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, hudi, mysql-server, tez-on-yarn
emr-5.27.1	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.27.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.26.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.25.0	2.3.5	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.24.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.24.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.23.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.23.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.22.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.21.2	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.21.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.21.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.20.1	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.20.0	2.3.4	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.19.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.19.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.18.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.18.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.17.2	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.17.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.17.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, emr-s3-select, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.16.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.16.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.15.1	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.15.0	2.3.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.14.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.14.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.14.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.13.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.13.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.12.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.12.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.12.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.4	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.11.3	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.2	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.11.1	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.11.0	2.3.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.10.1	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.10.0	2.3.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.9.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.9.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.8.3	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.2	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.8.1	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.8.0	2.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.7.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.7.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.6.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.6.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.5.4	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.3	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.5.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.5.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.5.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.4.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.4.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hive-hbase, hcatalog-server, hive-server2, mysql-server, tez-on-yarn
emr-5.3.2	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.3.1	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.3.0	2.1.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.2.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.2.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.2.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.1.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.1.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.0.3	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.2	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-5.0.1	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn
emr-5.0.0	2.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, hive-server, mysql-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.9.6	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.9.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.9.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.9.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.8.5	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.8.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.8.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.8.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.7.4	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.3	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.7.2	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.7.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.7.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, hive-server, mysql-server
emr-4.6.1	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.6.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.5.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.4.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.3.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.2.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server
emr-4.1.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

Amazon EMR 릴리스 레이블	Hive 버전	Hive와 함께 설치된 구성 요소
emr-4.0.0	1.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hive-metastore-server, hive-server, mysql-server

버전별 Hive 릴리스 정보

주제

- [Amazon EMR 6.8.0 - Hive 릴리스 정보]
- [Amazon EMR 6.7.0 - Hive 릴리스 정보]
- [Amazon EMR 6.6.0 - Hive 릴리스 정보]

Amazon EMR 7.8.0 - Hive 릴리스 정보

Amazon EMR 7.8.0 - Hive 변경 사항

유형	설명
버그 수정	CVE-2024-23953: Apache Hive: LLAP의 서명에 대한 타이밍 공격을 수정했습니다.

알려진 문제

- Amazon S3 Express One Zone을 출력 위치로 사용하는 Hive 삽입 덮어쓰기 쿼리의 경우 코어 사이트 구성을 `fs.s3a.directory.operations.purge.uploads`로 설정합니다`false`.

Amazon EMR 7.7.0 - Hive 릴리스 정보

Amazon EMR 7.7.0 - Hive 변경 사항

유형	설명
버그 수정	CVE-2024-29869: Apache Hive: 비제한적 권한으로 생성된 자격 증명 파일을 수정합니다.
버그 수정	Apache Ran에서 행 수준 필터링 정책이 활성화된 경우 SemanticException을 수정합니다.
버그 수정	LLAP 또는 ACID가 활성화된 경우 Tez 비동기 초기화 RR을 비활성화합니다.

알려진 문제

- Amazon S3 Express One Zone을 출력 위치로 사용하는 Hive 삽입 덮어쓰기 쿼리의 경우 코어 사이트 구성을 `fs.s3a.directory.operations.purge.uploads`로 설정합니다`false`.

Amazon EMR 7.6.0 - Hive 릴리스 정보

Amazon EMR 7.6.0 - Hive 변경 사항

유형	설명
개선 사항	ORC 비 ACID 파티셔닝 테이블에 대한 빠른 S3 접두사 나열 기능 추가
Feature	S3AFileSystem에서 Hive 쓰기 쿼리를 위한 매직 커미터에 대한 지원 추가

알려진 문제

- Amazon S3 Express One Zone을 출력 위치로 사용하는 Hive 삽입 덮어쓰기 쿼리의 경우 코어 사이트 구성을 `fs.s3a.directory.operations.purge.uploads`로 설정합니다`false`.

Amazon EMR 7.6.0 - 새 구성

분류	명칭	Default	설명
hive-site	`hive.exec.fast.s3.partition.discovery.enabled`	true	분할 계산에 빠른 S3 파티션 검색을 사용할지 여부입니다. 이렇게 하면 지원되는 파일 형식인 ORC에 대한 접두사 기반 목록이 활성화됩니다. 이 기능은 S3 Express One Zone 스토리지 클래스가 지원하지 않는 S3 API 파라미터를 사용합니다. 이 기능을 사용할 때는

분류	명칭	Default	설명
			이 기능을 비활성화합니다.
hive-site	hive.exec.fast.s3.partition.discovery.max.thread.threshold	128	빠른 S3 파티션 검색을 위한 최대 병렬 처리 수준입니다.
hive-site	hive.exec.fast.s3.partition.discovery.parallelism	10	빠른 S3 파티션 검색의 단일 실행의 병렬 처리 정도입니다. 이 구성은 이 로 hive.exec.fast.s3.partition.discovery.enabled 설정된 경우에만 적용됩니다. true

분류	명칭	Default	설명
hive-site	`hive.blobstore.output-committer.magic.track.commits.in.memory.enabled`	true	Hive를 사용하는 매직 커미터가 메모리에서 보류 중인 모든 커밋을 추적해야 합니까?를 전환하는 플래그 Magic 커미터에는 커밋 데이터를 메모리에 저장할 수 있는 옵션이 있어 S3 호출을 줄여 TaskCommit 작업 속도를 높일 수 있습니다. 이 구성은 하둡 구성을 재정의합니다. `fs.s3a.committer.magic.track.commits.in.memory.enabled`

분류	명칭	Default	설명
hive-site	hive.blobstore.output-committer.dp.skip.task.staging.dir.creation	true	전환할 플래그 매직 커미터가 blobstore에서 dp 스테이징 경로를 생성해야 합니까? 이 플래그는 Hive가를 통해 Magic Committer 를 사용할 때 메모리에서 커밋을 추적하는 경우에만 적용됩니다hive.blobstore.output-committer.magic.track.commits.in.memory.enabled . 기본적으로 true로 설정되지만가 hive.blobstore.output-committer.magic.track.commits.in.memory.enabled 활성화되어 있고 blobstore에 작업 생성 시도 경로의 추가 S3 호출을 저장하는 경우에만 적용됩니다.

분류	명칭	Default	설명
hive-site	`hive.blobstore.output-committer.magic.disable.fs.cache.for.llap`	true	Magic Committer를 사용할 때 LLAP의 쓰기 흐름에서 blobstore FS 캐시를 비활성화해야 하는지 여부를 전환하는 플래그입니다. LLAP가 활성화되면이 플래그가 표시되고 기본적으로 true로 설정됩니다.

Amazon EMR 7.5.0 - Hive 릴리스 정보

Amazon EMR 7.5.0 - Hive 변경 사항

유형	설명
개선 사항	HiveCLI에서 기존 세션을 사용하려고 하는 동안 Tez 세션이 열릴 때까지 기다리는 최대 시간을 10초로 늘렸습니다.
개선	LIMIT을 사용하여 간단한 선택 쿼리의 성능을 개선하도록 구성 튜닝

Amazon EMR 7.4.0 - Hive 릴리스 정보

Amazon EMR 7.4.0 - Hive 변경 사항

유형	설명
업그레이드	HIVE-28191: 하둡 버전을 3.4.0으로 업그레이드
업그레이드	hadoop 셰이딩된 protobuf를 3.21로 업그레이드
업그레이드	commons-cli를 1.5.0으로 업그레이드

유형	설명
업그레이드	commons-compress를 1.24.0으로 업그레이드
업그레이드	commons-io를 2.14.0으로 업그레이드
업그레이드	commons-lang3를 3.21.0으로 업그레이드
개선	HiveCLI에서 기존 세션을 사용하려고 하는 동안 Tez 세션이 열릴 때까지 기다리는 시간을 10초로 변경합니다.
개선	LIMIT을 사용하여 간단한 선택 쿼리에 대해 Tez DAG에서 단락 메커니즘 활성화
개선	HIVE-21100: TEZ 엔진 및 UNION 절을 사용할 때 발생한 테이블 하위 디렉터리의 평면화 허용
버그 수정	HIVE-25095: Beeline/hive -e 명령이 후행 따옴표가 있는 쿼리를 처리할 수 없음
버그 수정	HIVE-13781: 파티션 디렉터리가 존재하지 않을 때 FileNotFoundException에서 Tez 작업이 실패했습니다.
버그 수정	HIVE-28480: 이전 RS의 조인 브랜치 간에 파티션 해시 생성기 불일치 시 SMB 비활성화

Amazon EMR 7.4.0 - 새 구성

분류	명칭	Default	설명
hive-site	hive.ignore.failure.partition.dir.not.found	false	테이블 파티션이 있지만 실제 객체 스토리지 경로가 없는 경우 실패를 무시합니다.
hive-site	hive.tez.union.flatten.subdirectories	false	테이블에 데이터를 쓰고 UNION ALL이 쿼

분류	명칭	Default	설명
			리의 마지막 단계인 경우 Tez의 Hive는 UNION ALL의 각 브랜치에 대한 하위 디렉터리를 생성합니다. 이 속성이 활성화되면 하위 디렉터리가 제거되고 파일의 이름이 변경되어 상위 디렉터리로 이동합니다. hive.blobstore.use.output-committer가 활성화되어 있으면 효과가 없습니다.

Amazon EMR 7.3.0 - Hive 릴리스 정보

Amazon EMR 7.3.0 - Hive 변경 사항

유형	설명
Feature	HIVE-18728 - SSL로 webHCat을 보호합니다.
개선	LLAP 데몬 웹 UI에 대한 SSL 키 스토어 자격 증명 구성을 지원합니다.
개선	Hive 메타스토어 서버의 SSL 호스트 이름 확인을 제어하는 옵션을 제공합니다.
버그 수정	HIVE-26541 - WebHCat Service를 시작할 때 NPE입니다.
버그 수정	HIVE-23011 - 공동 작업 옵티마이저는 조인을 비교할 때 잔여 조건자를 확인해야 합니다.

유형	설명
버그 수정	javax.security.sasl.SaslException 수정: 전송 중 암호화가 활성화된 경우 HMS와 Namenode 간에 클라이언트와 서버 간에 공통 보호 계층이 없습니다.
버그 수정	orc 분할의 끝이 블록 위치의 시작과 겹치는 IOException을 수정합니다.
버그 수정	열 이름에 쉼표 문자가 포함되어 있고 CSVSerde를 사용하는 경우 항상 쉼표로 분할하는 대신 열 이름 구분 기호를 사용합니다.

Amazon EMR 7.3.0 - 새 구성

분류	명칭	Default	설명
hcatalog-webhcat-site	templeton.use.ssl	false	WebHCat 서버에 SSL 암호화를 사용하려면 이 값을 true로 설정합니다.
hcatalog-webhcat-site	templeton.keystore.path		WebHCat 서버의 SSL 인증서 키 스토어 위치입니다.
hcatalog-webhcat-site	templeton.keystore.password		WebHCat 서버의 SSL 인증서 키 스토어 암호입니다.
hcatalog-webhcat-site	templeton.ssl.protocol.blacklist	SSLv2, SSLv3	WebHCat 서버에 대해 비활성화할 SSL 버전입니다.

분류	명칭	Default	설명
hcatalog-webhcat-site	templeton.host	0.0.0.0	WebHCat 서버가 수신 대기할 호스트 주소입니다.
hive-site	hive.metastore.ssl.enable.hostname.verification	false	SSL/TLS 핸드셰이킹 중에 호스트 이름 확인을 제어합니다.
hive-site	hive.llap.daemon.web.ssl.keystore.path		LLAP 데몬 웹 UI의 SSL 인증서 키 스토어 위치입니다.
hive-site	hive.llap.daemon.web.ssl.keystore.password		LLAP 데몬 웹 UI의 SSL 인증서 키 스토어 암호입니다.
hive-site	hive.metastore.hadoop.rpc.protection.override.to.authentication	false	활성화하면 HMS는 항상 구성 세트에서 인증을 위한 hadoop.rpc.protection 값을 재정의합니다.

Amazon EMR 7.2.0 - Hive 릴리스 정보

Amazon EMR 7.2.0 - Hive 변경 사항

유형	설명
업그레이드	Parquet 1.13.1 - Parquet이 113.1로 업그레이드되었습니다.
개선	HIVE-12930 - LLAP용 SSL 셔플을 지원합니다.
개선	HIVE-23062 - 사용자 지정 Tez 대기열의 TLS 및 Yarn HA 모드에서 Yarn RM URL을 확인하는 Hive입니다.

유형	설명
버그 수정	HIVE-27952 - KeyStore에 여러 인증서가 있는 경우 Hive가 SslContextFactory를 생성하지 못합니다.
버그 수정	HIVE-28085 - YarnQueueHelper가 HTTPS 지원 YARN WebService에 액세스하지 못합니다.
버그 수정	HIVE-26436 - initializeOp가 호출되지 않고 닫히는 경우 MR NullPointerException의 Hive입니다. 연산자가 초기화되지 않은 경우 닫기를 건너뜁니다.

Amazon EMR 7.2.0 - 새 구성

분류	명칭	Default	설명
hive-site	hive.llap.shuffle.ssl.enabled	false	LLAP용 SSL 셔플을 활성화하려면 tez.runtime.shuffle.ssl.enable과 함께 true로 설정합니다.

Amazon EMR 7.1.0 - Hive 릴리스 정보

Amazon EMR 7.1.0 - Hive 변경 사항

유형	설명
버그 수정	HIVE-24381 - 헤더/바닥글 건너뛰기가 포함된 경우 압축 텍스트 입력은 0개의 행을 반환합니다.

유형	설명
버그 수정	HIVE-24190 – LLAP: ShuffleHandler는 TEZ-4233에 따라 DISK_ERROR_EXCEPTION 을 반환할 수 있습니다.
버그 수정	HIVE-23073 – Shade Netty.
버그 수정	HIVE-23073 – Shade Netty 및 netty 4.1.48.Final 로 업그레이드.
버그 수정	HIVE-23148 - Netty 셰이딩으로 Llap 외부 클라이언트 흐름이 중단되었습니다.
버그 수정	HIVE-25180 - Netty를 업그레이드합니다.
버그 수정	HIVE-24524 – LLAP ShuffleHandler: Netty4로 업그레이드하고 가능한 경우 Hive에서 Netty3 종속 항목을 제거합니다.
버그 수정	HIVE-28000 - Hive QL: 유형 강제 변환을 수행할 수 없는 경우 'not in' 절에서 잘못된 결과를 제공합니다.
버그 수정	HIVE-27993 – Netty4 ShuffleHandler는 보스 스레드 1개를 사용해야 합니다.
업그레이드	Netty를 4.1.100.Final로 업그레이드
업그레이드	Jetty를 9.4.53.v20231009로 업그레이드
업그레이드	Zookeeper를 3.9.1로 업그레이드

Amazon EMR 7.1.0 - Hive 변경 사항

- Amazon EMR 7.1은 Hive를 Netty 4.1.100.Final로 업그레이드하여 Netty3의 보안 취약성을 해결합니다. Hive-druid-handler는 netty3에 종속되어 있으므로 Hive는 Amazon EMR 7.1에서 Hive 클래스 경로에 `hive-druid-handler` JAR을 포함하지 않습니다. Druid 핸들러가 Netty의 4.1.100.Final 이

상 버전을 지원하면 향후 Amazon EMR 릴리스가 Hive의 클래스 경로에 이를 포함합니다. Amazon EMR 릴리스 7.1 이상에서 `hive-druid-handler` JAR이 필요한 경우 AWS 에 문의하십시오.

Amazon EMR 7.0.0 - Hive 릴리스 정보

Amazon EMR 7.0.0 - Hive 변경 사항

유형	설명
업그레이드	Hive 런타임은 이제 기본적으로 Java 17을 사용합니다. 자세한 내용은 EMR 7.0.0 릴리스 가이드를 참조하세요.
백포트	HIVE-17709: sun.misc.Cleaner 참조 제거
버그 수정	LLAP 또는 ACID가 활성화된 경우 Tez 비동기 초기화 RR 비활성화

Amazon EMR 6.15.0 - Hive 릴리스 정보

Amazon EMR 6.15.0 - Hive 변경

유형	설명
Feature	TEZ-4397 지원 – Tez 비동기 스플릿 오프닝의 경우 Hive에서는 이제 Tez 비동기 분할 열기에 설명된 Tez 구성을 지원합니다.
버그 수정	HIVE-25400 – `BytesColumnVector` 에 있는 오프셋 업데이트를 `setValPreallocated` 로 이동합니다.
버그 수정	HIVE-25190 – `BytesColumnVector` 에서 여러 개의 작은 할당을 수정합니다.
버그 수정	작업자 노드에서 LLapDaemon을 시작할 때 NoClassDefFound 예외를 방지하기 위해 llap 서버로 빈 모듈을 패키징합니다.

유형	설명
업그레이드	Apache Hadoop을 3.3.6으로 업그레이드합니다.
업그레이드	HIVE-26684 – `maven-shade-plugin` 을 3.4.1로 업그레이드합니다.
개선	Amazon EMR 클러스터 시작 시간을 줄이려면 HCatalog 시작 스크립트에서 15초의 휴면 시간을 제거하세요.

Amazon EMR 6.14.0 - Hive 릴리스 정보

Amazon EMR 6.14.0 - Hive 변경

유형	설명
개선 사항	HIVE-26762: `HiveFilterSetOpTransposeRule` 에서 피연산자 정리 제거
버그 수정	HIVE-27582: FetchOperator에서 HBase 테이블 입력 형식을 캐시하지 않음
버그 수정	HIVE-26452: JOIN를 MAPJOIN로 변환하는 중 NPE, JOIN 열이 두 번 이상 참조됨
버그 수정	HIVE-26416: `AcidUtils.isRawFormatFile()` 에서 non-ORC 파일에 대해 `InvalidProtocolBufferException` 예외 발생
버그 수정	HIVE-26105: 설명 열에 특정 중국어 문자가 포함된 경우 열 표시에서 추가 값을 표시함
버그 수정	HIVE-25864: Hive 쿼리 최적화로 인해 윈도우 함수를 포함하는 조건부 푸시다운에 대해 잘못된 계획이 생성됨

유형	설명
버그 수정	HIVE-25224: 서로 다른 `bucketing_versions` 의 테이블을 포함하는 여러 INSERT 문에서 오류 발생
버그 수정	HIVE-24151: 문자열에 non-ASCII 문자가 포함된 경우 `MultiDelimitSerDe` 에서 이동시킴
버그 수정	HIVE-23606: (LLAP) `EncodedReaderImpl` 에 대해 `DirectByteBuffer` 정리 지연
버그 수정	HIVE-22165: `SessionManager.closeSession` 에서 HIVE-14296을 통해 도입된 동기화로 인해 사용량이 많은 Hive 서버에서 지연 시간이 길어짐
버그 수정	HIVE-21304: 버킷팅 버전 사용 기능이 개선됨

Amazon EMR 6.13.0 - Hive 릴리스 정보

Amazon EMR 6.13.0 - Hive 변경

유형	설명
개선 사항	Python3를 지원하도록 Python 스크립트 업그레이드
개선 사항	HIVE-27097: MetaStore 클라이언트 및 서버의 재시도 전략 개선
버그 수정	HIVE-21778: CBO: null 허용은 항상 쿼리에서 필터를 누락시키므로 'Struct is not null'이 평가됨
버그 수정	HIVE-21009: 사용자가 바인드 사용자를 설정하는 기능 추가

유형	설명
버그 수정	HIVE-22661: 경로에 데이터가 로드된, 버킷팅되지 않은 테이블에서 압축 실패
버그 수정	HIVE-19718: 파티션을 대량으로 추가하면 각 파티션의 테이블도 가져옴
버그 수정	HIVE-22173: 여러 측면 보기가 있는 쿼리가 컴파일 중에 중단됨
버그 수정	HIVE-27088: 포스트 조인 필터가 있는 내부 및 외부 조인을 병합하면 잘못된 결과가 생성됨
버그 수정	HIVE-21935: Hive 벡터화: UDF 벡터화 시 성능이 저하됨
버그 수정	HIVE-25299: UTC 외 시간대에서 타임스탬프를 숫자 데이터 유형으로 캐스팅하는 작업이 올바르지 않음
버그 수정	HIVE-24626: LLAP: 모든 IO 엘리베이터 스레드가 사용 중이어서 대기열이 가득 찬 다른 리더를 대기열에 넣는 경우 리더 스레드가 모두 소진될 수 있음
버그 수정	HIVE-27029: 파일 시스템 닫힘 오류로 Hive 쿼리 실패, HIVE-26352에서 수정 완료
버그 수정	HIVE-26352: 압축 시 GSS 예외로 인해 Tez 대기열 액세스 확인 실패
버그 수정	HIVE-24590: 작업 로깅에서 여전히 log4j 어펜더가 누출됨
버그 수정	HIVE-24552: loadDynamicPartitions에서 HMS 연결이 누출되거나 누적될 수 있음
버그 수정	HIVE-27069: 버킷 맵 조인에서 잘못된 결과

유형	설명
버그 수정	HIVE-27344: RecordReaderImpl#close에 널 검사 추가
버그 수정	HIVE-27439: 소수점에서 공백 지원
버그 수정	HIVE-27267: 하위 쿼리가 있는 버킷팅된 10진수 열에서 버킷 맵 조인을 수행할 때 잘못된 결과가 생성됨
버그 수정	HIVE-21986: HiveServer 웹 UI: 기본 응답 헤더에 Strict-Transport-Security 설정
버그 수정	HIVE-22148: S3A 위임 토큰이 압축기의 작업 구성에 추가되지 않습니다.
버그 수정	HIVE-22622: Hive에서 속성 이름이 중복된 구문을 생성할 수 있음
버그 수정	HIVE-22008: LIKE 연산자는 여러 줄 입력과 일치해야 함
버그 수정	HIVE-23144: LLAP: serviceStop에서 QueryTracker 정리 허용
버그 수정	HIVE-22391: Hive 쿼리 결과 캐시를 확인하는 중 NPE
버그 수정	HIVE-23305: 경합 조건으로 인해 LlapTaskSchedulerService addNode에서 NullPointerException 발생
버그 수정	HIVE-22178: Parquet FilterPredicate에서 SchemaEvolution 이후 CastException 발생
버그 수정	HIVE-21517: AggregateStatsCache 수정
버그 수정	HIVE-21825: 액티브 및 패시브 HA가 활성화된 경우 클라이언트 오류 메시지 개선

유형	설명
버그 수정	HIVE-23389: FilterMergeRule에서 Assertion Error가 발생할 수 있음
버그 수정	HIVE-22767: Beeline에서 설명의 세미콜론을 제대로 구문 분석하지 않음
버그 수정	HIVE-22996: BasicStats 구문 분석에서 null 또는 빈 문자열을 사전에 확인해야 함
버그 수정	HIVE-22808: HiveRelFieldTrimmer에서 HiveTableFunctionScan을 처리하지 않음
버그 수정	HIVE-22437: 메타데이터를 잠글 때 LLAP 메타데이터 캐시 NPE
버그 수정	HIVE-22606: AVroserde에서 정보 수준 아래 avro.schema.literal을 로깅함
버그 수정	HIVE-22713: Join-Fil(*)-RS 구문에 대해 지속적 전파를 수행해서는 안 됨
버그 수정	HIVE-21624: LLAP: 스레드 수준의 CPU 지표가 손상됨
버그 수정	HIVE-22815: MROutput에서 불필요한 파일 시스템 객체 생성 감소
버그 수정	HIVE-23060: 'Grouping sets expression is not in GROUP BY key. Error encountered near token' 오류로 쿼리 실패
버그 수정	HIVE-22236: `NOT IN` 하위 쿼리가 포함된 보기를 선택할 때 보기 생성 실패
버그 수정	HIVE-19886: -hiveconf hive.log.file을 사용하는 경우 로그가 두 개의 파일로 전달될 수 있음

유형	설명
버그 수정	HIVE-20620: 동적 파티셔닝을 사용하여 버킷팅 및 정렬된 MM 테이블에 삽입할 때 매니페스트 충돌
버그 수정	HIVE-14557: SkewJoin 및 Mapjoin이 모두 활성화된 경우 Nullpointer
버그 수정	HIVE-20471: 기본 데이터베이스 경로를 가져오는 중 문제 발생
버그 수정	HIVE-20598: HiveAlgorithmsUtil 계산에서 오타 수정
버그 수정	HIVE-14737: Kerberos 지원 Hive Server 2 웹 UI의 /logs에 액세스하는 중 오류 발생
버그 수정	HIVE-20733: GenericUDFOPEqualNS 는 계획 설명에 =를 사용할 수 없음
버그 수정	HIVE-20848: UpdateInputAccessTimeHook 설정 후 테이블을 찾을 수 없음 오류로 쿼리에 실패합니다.
버그 수정	HIVE-18929: HiveStringUtils.java에서 humanReadableInt 메서드에 경합 조건이 있습니다.
버그 수정	HIVE-20841: LLAP: 동적 포트를 구성 가능하게 설정
버그 수정	HIVE-20930: 필터 모드의 VectorCoalesce가 적용되지 않음
버그 수정	HIVE-21007: 세미조인 + 집합으로 잘못된 계획이 생성될 수 있음

유형	설명
버그 수정	HIVE-21074: IS NOT NULL 조건에서 Hive 버킷팅된 테이블 쿼리 정리가 작동하지 않음
버그 수정	HIVE-21223: 파티션이 존재하지 않는 경우 CachedStore가 null 파티션을 반환함
버그 수정	HIVE-19625: Hive#copyFiles에서 잠재적 NPE 및 실제 예외 숨기기
버그 수정	HIVE-17020: 적극적인 RS 중복 제거로 OP 트리 브랜치가 잘못 제거될 수 있음
버그 수정	HIVE-20168: ReduceSinkOperator 로깅 숨김
버그 수정	HIVE-20879: 프로젝션 표현식에서 null을 사용하면 CastException이 발생함
버그 수정	HIVE-20888: TxnHandler: 변경 불가 목록에서 sort() 직접 호출됨
버그 수정	HIVE-19948: 문자열 안에 따옴표가 있는 경우 HiveCli가 명령을 세미콜론으로 제대로 분할하지 않음
버그 수정	HIVE-20621: resultset.next에서 GetOperationStatus가 직접 호출되어 증분식으로 속도가 느려짐
버그 수정	HIVE-20854: 합리적인 기본값: Hive의 Zookeeper 하트비트 간격이 20분임, 2로 변경
버그 수정	HIVE-20330: HCatLoader는 입력이 여러 개인 작업에 대해 여러 InputJobInfo 객체를 처리할 수 없음

유형	설명
버그 수정	HIVE-20787: MapJoinBytesTableContainer dummyRow 케이스에서 재사용을 처리하지 않음
버그 수정	HIVE-20331: union all, lateral view, Join을 포함한 쿼리가 'cannot find parent in the child operator로 실패
버그 수정	HIVE-19968: UDF 예외가 처리되지 않음
버그 수정	HIVE-20410: 트랜잭션 테이블에 대한 삽입 덮어쓰기가 중단되면 'Not enough history available for…' 오류 발생
버그 수정	HIVE-20059: Hive 스트리밍은 예외 발생 시 무조건 접두사 음영 처리를 시도해야 함
버그 수정	HIVE-19424: MetaDataFormatters에서 NPE
버그 수정	HIVE-20355: HiveConnection.setSchema의 정리 파라미터
버그 수정	HIVE-20858: Utilities.createEmptyBuckets의 구성에서 Serializer가 올바르게 시작되지 않음
버그 수정	HIVE-20424: schematool이 beeline 기록을 오염시키지 않아야 함
버그 수정	HIVE-20338: LLAP: OSIX 변형 의미 체계와 함께 HDFS 프로토콜 구현을 포함하는 파일 시스템에서 합성 file-id 강제 실행
버그 수정	HIVE-11708: 논리 연산자는 NULL과 함께 ClassCastExceptions 생성
버그 수정	HIVE-21082: HPL 및 SQL에서 선언문은 문자 유형의 변수를 지원하지 않음

유형	설명
버그 수정	HIVE-16690: LLAP 클러스터 크기를 기반으로 Tez Cartesian 제품 엣지 구성
버그 수정	HIVE-14516: OrcInputFormat.SplitGenerator.callInternal
버그 수정	HIVE-20981: streaming, AbstractRecordWriter에서 HeapMemoryMonitor 누출
버그 수정	HIVE-20043: HiveServer2: SessionState에서 AtomicBoolean 주변에 정적 동기화 블록이 있음
버그 수정	HIVE-20191: 패치가 비어 있어도 PreCommit 패치 애플리케이션이 실패하지 않음
버그 수정	HIVE-20400: create table은 잠재적인 FS 모호성을 방지하려면 항상 완전한 경로를 사용해야 함
버그 수정	스큐된 열에 액세스하기 전에 skewedInfo에 대한 null 검사 추가

Amazon EMR 6.12.0 - Hive 릴리스 정보

Amazon EMR 6.12.0 - Hive 변경

유형	설명
개선 사항	JDK 11 및 JDK 17 런타임에 대한 지원 추가됨
개선 사항	S3 Select를 사용할 때 대소문자를 구분하고 예약된 키워드 열 이름을 쿼리하기 위한 지원을 추가합니다. 이를 사용하려면 's3select.column.mapping' = "`column1:fieldName1 , column2:fieldName2 ,...`" 형식으로 테이블 속성 정의

유형	설명
개선 사항	HIVE-23133: 숫자 연산은 하드웨어 아키텍처에 따라 결과가 다를 수 있음
개선 사항	HIVE-27145: HIVE-23133의 후속 조치로 나머지 산술 함수에 대해 StrictMath 사용
버그 수정	EMR Hive 6.4.0에서 HIVE-22900을 이식하여 발생한 get_partitions_by_filter 및 get_num_partitions_by_filter HMS API의 와일드카드 비호환성 수정
버그 수정	HIVE-26736: WITH 절이 포함된 중첩된 보기에 대한 권한 부여 실패
버그 수정	HIVE-22416: 병렬 실행이 활성화된 경우 MR 관련 작업 로그가 누락됨
버그 수정	HIVE-19653: 그룹화 세트를 사용하는 groupby의 잘못된 조건자 푸시다운
버그 수정	HIVE-22094: ClassCastException으로 쿼리 실패: hive.ql.exec.vector.DecimalColumnVector를 hive.ql.exec.vector.Decimal64ColumnVector로 캐스팅할 수 없음
버그 수정	HIVE-26340: 쿼리에 대문자 기간 함수가 있는 경우 벡터화된 PTF 연산자 실패
버그 수정	HIVE-26184: 일부 키가 심하게 스큐된 경우 GROUP BY를 포함하는 COLLECT_SET 속도가 매우 느려짐
버그 수정	HIVE-26373: Avro 데이터를 포함하는 HBase 테이블에서 타임스탬프를 읽을 때 ClassCastException 발생

유형	설명
버그 수정	HIVE-26388: CTAS 쿼리의 소스 테이블에 문자열 유형이 아닌 열이 있는 경우 ClassCast Exception 발생 업그레이드 HIVE-26172: Hive - CVE-2021-36373 및 CVE-2021-36374 때문에 Ant를 1.10.11로 업그레이드
버그 수정	HIVE-26114: 접두사 공백을 포함하는 dfs 명령을 사용하는 jdbc 연결 hiveserver2에서 발생하는 예외 수정
버그 수정	HIVE-26396: trunc 함수에 정밀도 인터셉션 관련 문제가 있으며, 결과에 0이 너무 많음
버그 수정	HIVE-26446: HiveProtologgingHook에서 파티셔닝된 테이블의 TablesWritten 필드를 채우지 못했습니다.
버그 수정	HIVE-26639: ConstantVectorExpression 및 ExplainTask는 기본 문자 세트에 의존할 수 없음
버그 수정	HIVE-22670: parquet 파일을 읽을 때 벡터화된 리더가 사용되는 경우 ArrayIndexOutOfBoundsException 발생
버그 수정	HIVE-23607: 권한 문제: 다른 보기에서 보기 생성에 성공하지만 보기 대체 실패
버그 수정	HIVE-25498: 31개가 넘는 count distinct 함수를 포함하는 쿼리가 잘못된 결과를 반환함
버그 수정	HIVE-25780: DistinctExpansion에서는 64개가 넘는 그룹화 세트 II를 생성함
버그 수정	HIVE-23868: 기간 함수 사양: 선행 0 및 후속 0 지원

유형	설명
버그 수정	HIVE-24539: ORInputFormat 스키마 생성은 열 구분 기호를 인식해야 함
버그 수정	HIVE-23476: LLAP: mmap 케이스에도 아레나 사전 할당
버그 수정	HIVE-25806: LlapCacheAwareFs에서 잠재적 누출 - Parquet, LLAP IO
버그 수정	HIVE-23498: ThriftHttpCliService에서 HTTP Trace 메서드 비활성화
버그 수정	HIVE-25729: 전체 초기화된 경우 ThriftUnionObjectInspector에 알려야 함
버그 수정	HIVE-23846: bitvectors의 불필요한 직렬화 및 역직렬화 방지
버그 수정	HIVE-24233: cbo가 비활성화된 상태로 except 하위 쿼리에서 nullpointer 발생
버그 수정	HIVE-24276: HiveServer2 loggerconf jsp 크로스 사이트 스크립팅(XSS) 취약성
버그 수정	HIVE-25721: 외부 조인 결과가 잘못됨
버그 수정	HIVE-25223: 제한이 있는 Select에서 기본 테이블이 아닌 테이블에 행을 반환하지 않음
버그 수정	HIVE-25794: CombineHiveRecordReader: 루프의 log 문으로 인해 메모리 부담 생성
버그 수정	HIVE-23602: 작업 핸들 세트에 Java 동시 패키지 사용
버그 수정	HIVE-24045: 기본 데이터베이스 생성 시기와 관련된 로깅 없음

유형	설명
버그 수정	HIVE-24305: 값을 따옴표로 묶으면 avro 십진수 스키마가 규모 및 정밀도를 올바르게 채우지 않음
버그 수정	HIVE-25040: 데이터베이스 케스케이드 삭제에서 영구 함수를 제거할 수 없음
버그 수정	HIVE-23501: 복합 유형을 기본 유형으로 변환할 때 VectorDeserializeRow에서 AOOB 발생
버그 수정	HIVE-23704: Thrift HTTP 서버에서 인증 핸들을 올바르게 처리하지 않음
버그 수정	HIVE-23529: row_deserialize에서 uniontype에 대해 CTAS 손상됨
버그 수정	HIVE-24144: HiveDatabaseMetaData의 getIdentifierQuoteString에서 잘못된 값이 반환됨
버그 수정	HIVE-23850: 제목이 그룹화 세트가 있는 열이 아닌 경우 PPD 허용
버그 수정	HIVE-25919: HBaseStorageHandler에서 부울 열 조건부를 푸시할 때 ClassCastException 발생
버그 수정	HIVE-25261: RetryingHMSHandler에서 대상에 대한 간단한 설명을 포함하는 MetaException을 줄 바꿈해야 함
버그 수정	HIVE-24792: 작업 중 잠재적 스레드 누출
버그 수정	HIVE-23409: 타임라인 서비스 중단으로 인해 TezSession 애플리케이션을 다시 열지 못하는 경우 재시도 후 SessionPool의 기본 TezSession이 닫힘

유형	설명
버그 수정	HIVE-23615: Beeline 명령 클래스에서 null 포인터 참조 해제하지 않음
버그 수정	HIVE-24849: 위치에 파일이 많은 경우 외부 테이블 소켓 제한 시간(3.1.2에 영향을 미침)
버그 수정	HIVE-25209: SUM 함수를 포함하는 SELECT 쿼리에서 예상치 못한 결과 생성
버그 수정	HIVE-23666: groupby 연산자에 그룹화 세트가 없으면 checkHashModeEfficiency 건너뜀
버그 수정	HIVE-23873: CBO가 꺼져 있는 경우 Hive JDBCStorageHandler 테이블 쿼리가 NPE로 실패함
버그 수정	HIVE-24149: HiveStreamingConnection에서 HMS 연결을 닫지 못함
버그 수정	HIVE-25561: 강제 종료된 작업에서 파일을 커밋할 수 없습니다(2.x 및 3.x 버전에 영향을 미침).
버그 수정	HIVE-25683: AcidUtils.isRawFormatFile에서 리더 닫기
버그 수정	HIVE-24294: TezSessionPool 세션에서 AssertionError 오류가 발생할 수 있음
버그 수정	HIVE-24182: 영구 UDF 관련 Ranger 권한 부여 문제
버그 수정	HIVE-22805: 조건부 배열 또는 맵을 사용한 벡터화가 구현되지 않아 오류 발생
버그 수정	HIVE-22828: Decimal64: NVL 및 CASE 문이 decimal64를 128로 암시적으로 변환

유형	설명
버그 수정	HIVE-21398: 추정 통계를 포함하는 열을 고유 키로 간주해서는 안 됨
버그 수정	HIVE-22490: 경로에 특수 문자가 포함된 jar를 추가하면 오류 발생
버그 수정	HIVE-22700: 권한이 없는 경우 압축 시 메모리 누수가 발생할 수 있음
버그 수정	HIVE-22053: 함수 생성 시 함수 이름이 정규화되지 않음
버그 수정	HIVE-22595: 외부 스키마가 있는 Avro 테이블 테이블에서 동적 파티션 삽입 실패
버그 수정	HIVE-21795: 파티션셔닝된 테이블에서 mapjoin 수행 시 롤업 요약 행이 누락될 수 있음
버그 수정	HIVE-22987: DataTypePhysicalVariation이 null인 경우 VectorCoalesce에서 ClassCast Exception 발생
버그 수정	HIVE-22219: 노드 관리자를 중단하여 LLAP 서비스 다시 시작 차단
버그 수정	HIVE-21793: hive.stats.fetch.column.stats가 false로 설정되었어도 CBO가 열 통계를 검색함
버그 수정	HIVE-22163: CBO: CBO를 활성화하면 추정이 비활성화된 경우에도 통계 추정이 켜짐
버그 수정	HIVE-18735: 유사 테이블 생성 시 트랜잭션 속성 손실

유형	설명
버그 수정	HIVE-22433: Hive JDBC 스토리지 핸들러: JDBC 데이터 소스의 BOOLEAN 및 TIMESTAMP DataType에서 잘못된 결과를 가져옴
버그 수정	HIVE-19430: 보류 중인 많은 이벤트에서 ObjectStore.cleanNotificationEvents OutOfMemory
버그 수정	HIVE-20785: JDBC DatabaseMetaData.getPrimaryKeys 메서드에서 잘못된 키 이름
버그 수정	HIVE-16116: beeline.properties에서 beeline.hiveconfvariables={}인 경우 Beeline에서 NPE 발생
버그 수정	HIVE-20066: hive.load.data.owner를 전체 보안 주체와 비교함
버그 수정	HIVE-20489: 쿼리 계획 설명 중지
버그 수정	HIVE-21033: 작업 닫기를 잊으면 추가로 HiveServer2 출력 잘림
버그 수정	HIVE-19888: SessionState의 잘못된 'METASTORE_FILTER_HOOK will be ignored' 경고
버그 수정	HIVE-20303: INSERT OVERWRITE TABLE db.table PARTITION (...) IF NOT EXISTS에서 InvalidTableException 발생
버그 수정	HIVE-16144: CompactionInfo에 equals/hashCode가 있지만 Set에서 사용됨

유형	설명
버그 수정	HIVE-20818: WHERE 하위 쿼리로 생성된 보기는 하위 쿼리에서 참조된 보기를 직접 입력으로 간주함
버그 수정	HIVE-21005: LLAP: 분할당 스트라이프를 더 많이 읽으면 ZlibCodecs에서 누출 발생
버그 수정	HIVE-20771: 빈 구문에서 LazyBinarySerDe에 실패합니다.
버그 수정	HIVE-18852: alter table 검증에서 잘못된 오류 메시지
버그 수정	HIVE-21124: HPL/SQL은 CREATE TABLE LIKE 문을 지원하지 않음
버그 수정	HIVE-20935: EC2 에서 llap 패키지 tarball 업로드에 실패하여 LLAP 서비스 시작 실패
버그 수정	HIVE-20409: Hive ACID: 업데이트, 삭제, 병합 시 hdfs 스테이징 디렉터리가 정리되지 않음
버그 수정	HIVE-20570: hive.optimize.union.remove=true 가 설정된 Union ALL에서 잘못된 계획 포함
버그 수정	HIVE-20421: hive-default.xml.template의 '\b' 엔터티에 잘못된 문자가 있음
버그 수정	HIVE-19133: HS2 WebUI 단계별 성능 지표가 올바르게 표시되지 않음
버그 수정	HIVE-18977: 파티션을 나열하면 JDO 및 직접 SQL과 다른 결과 반환
버그 수정	HIVE-20034: 이전 버전과의 호환성을 위해 메타스토어 예외 처리 변경 사항 롤백

유형	설명
버그 수정	HIVE-20672: LlapTaskSchedulerService의 로깅 스레드에서 수정된 모든 간격을 보고해야 함
버그 수정	HIVE-12812: 집계 함수를 포함하는 집합을 지원하도록 기본적으로 mapred.input.dir.recursive 활성화
버그 수정	HIVE-20147: Hive 스트리밍 모으기가 동기화된 로깅에 충족됨
버그 수정	HIVE-19203: HiveMetaStore에서 스레드 안전 문제
버그 수정	HIVE-20091: Tez: FileSinkOperator 출력에 대한 보안 인증 추가
버그 수정	HIVE-16906: Hive ATSHook는 ATS에 연결하기 전에 yarn.timeline-service.enabled를 검사해야 함
버그 수정	HIVE-20714: 단일 속성에 대한 SHOW tblproperties에서 이름 열에 값을 반환함
버그 수정	HIVE-24730: Shims 클래스는 hive-site.xml 및 tez-site.xml의 값을 자동으로 재정의함
버그 수정	HIVE-22055: 텍스트 파일에서 데이터를 로드한 후 select count에서 잘못된 결과 제공

Amazon EMR 6.11.0 - Hive 릴리스 정보

Amazon EMR 6.11.0 - Hive 변경

유형	설명
개선 사항	파티션 삭제 성능을 향상시키기 위해 파티션의 멀티스레드 삭제 지원이 추가됨
개선 사항	인코딩된 Hive 쿼리 파일 읽기 지원
개선 사항	Tez 기반 Hive 작업에 대해 기본적으로 Tez 셔플 핸들러를 활성화함
버그	hive.groupby.skewindata가 활성화된 경우 결정론적 키 분배를 활성화하는 옵션을 reducer에 추가하여 잘못된 결과 수정(HIVE-20220에 보고됨)
버그	기본 파티션 이름을 구성할 때 통계 계산 실패를 수정함
버그	전송 중 암호화가 활성화된 클러스터에서 HiveServer2에 대해 기본적으로 SSL을 구성할 때 전달되는 모든 사용자 지정 SSL 분류 파라미터 인식
백포트	HIVE-23617: storage-api FindBug 문제를 수정함
백포트	HIVE-26408: 벡터화: 스크래치 열의 할당 해제를 수정하고 하위 ConstantVectorExpression을 출력으로 재사용하지 않습니다.
백포트	HIVE-23614: removeTempOrDuplicateFiles를 위해 항상 HiveConfig 전달
백포트	HIVE-23354: compareTempOrDuplicateFiles에서 파일 크기 안전성 검사 제거

유형	설명
백포트	HIVE-20344: AccessControlException 예외가 발생하던 SBA에 대한 PrivilegeSynchronizer 문제를 수정했습니다. 또한 권한 동기화 도구를 비활성화하기 위해 hive.privilege.synchronizer 속성을 도입함
백포트	HIVE-15826: 모든 SerDes에 대해 'serialization.encoding' 구성 지원
백포트	HIVE-18284: dynpart 정렬 최적화에서 'distribute by' 절을 사용하여 데이터를 삽입할 때 NPE 수정
백포트	HIVE-24930: 하위 연산자에서 Operator.setDone() short-circuit이 벡터화된 코드 경로에서 사용되지 않음(childSize == 1인 경우)
백포트	HIVE-24523: LazySimpleSerde의 벡터화된 읽기 경로가 타임스탬프에 대해 SERDEPROPERTIES를 인식하지 않음
백포트	HIVE-23265: 제한 및 오프셋이 설정된 경우 중복된 행 세트가 반환됨
백포트	HIVE-21492: VectorizedParquetRecordReader가 thrift 및 사용자 지정 도구를 사용하여 생성된 parquet 파일을 읽을 수 없음
백포트	HIVE-22540: 벡터화: Decimal64 열이 VectorizedBatchUtil.makeLikeColumnVector()에서 작동하지 않음
백포트	HIVE-22588: 벡터 groupby 모드를 전환할 때 나머지 그룹화 세트의 나머지 행 비우기
백포트	HIVE-22551: BytesColumnVector initBuffer에서 벡터 및 길이를 일관되게 정리해야 함

유형	설명
백포트	HIVE-22448: CBO: group-by 키를 사용하여 여러 count distinct 확장
백포트	HIVE-22248: 통계가 지속되는 문제 수정
백포트	HIVE-22210: 벡터화는 필터링과 관련된 계산 출력 열을 재사용할 수 있음
백포트	HIVE-21531: 벡터화: 모든 NULL 해시코드가 Murmur3을 사용하여 계산되지 않음
백포트	HIVE-20419: 벡터화: 해시맵 키에 사용된 후 VectorPartitionDesc 변형 방지
백포트	HIVE-19388: VectorMapJoinCommonOperator 초기화 중 ClassCastException 발생
백포트	HIVE-21584: Java 11 준비: 시스템 클래스 로더는 URLClassLoader가 아님
백포트	HIVE-25107: 클래스 경로 로깅은 디버그 수준이어야 함(2271)
백포트	HIVE-22097: Java 11에서 호환되지 않는 java.util.ArrayList
백포트	HIVE-23938: LLAP: JDK11 - 일부 GC 로그 파일 로테이션 관련 jvm 인수를 더 이상 사용할 수 없음
백포트	HIVE-26226: upgrade-acid의 hive-metastore에서 jdk.tools dep 제외
백포트	HIVE-17879: Datanucleus Maven 플러그인 업그레이드

유형	설명
백포트	HIVE-27004: DateTimeFormatterBuilder#appendZoneText는 8 이상의 Java 버전에서 'UTC+'를 구문 분석할 수 없음
백포트	HIVE-16812: VectorizedOrcAcidRowBatchReader가 삭제 이벤트를 필터링하지 않음
백포트	HIVE-17917: VectorizedOrcAcidRowBatchReader.computeOffsetAndBucket 최적화
백포트	HIVE-19985: ACID: 읽기 전용 쿼리의 ROW__ID 섹션 디코딩 건너뛰기
백포트	HIVE-20635: VectorizedOrcAcidRowBatchReader가 원래 파일에 대한 삭제 이벤트를 필터링하지 않음
업그레이드	Javadocfmf 3.3.1로 업그레이드
업그레이드	Javassist를 3.24.1-GA로 업그레이드
업그레이드	apache-directory-server를 2.0.0-M14로 업데이트

새 구성

명칭	분류	설명
hive.metastore.fs.drop.partition.threads	hive-site	파티션 삭제 스레드 풀의 코어 스레드 수
hive.metastore.fs.drop.partition.keepalive.time	hive-site	스레드 풀에서 유휴 파티션 삭제 비동기 스레드가 종료 전 새 작업이 도착하길 기다리는 시간(초).

명칭	분류	설명
hive.metastore.fs.drop.partition.threadpool.max.queue.size	hive-site	파일 시스템에서 파티션을 삭제하기 위해 스레드 풀에서 사용할 최대 대기열 크기.
hive.groupby.enable.deterministic.distribution	hive-site	reducer에 대한 결정론적 키 분배를 활성화합니다. 무작위 파티셔닝에 사용되는 rand 함수를 직접 호출하는 동안 일정한 시드 값을 전달합니다.
hive.privilege.synchronizer	hive-site	HiveServer2에서 외부 권한 부여자의 권한을 정기적으로 동기화할지 여부.
hive.cli.query.file.encoding	hive-site	cli 인수에 제공된 모든 유형의 쿼리 파일(query 파일, init 쿼리 파일, rc 파일 등)에 대한 파일 인코딩
hive.emr.tez.shuffle.enabled	hive-site	Tez 기반 Hive는 이제 기본 셔플 핸들러로 mapreduce_shuffle 대신 tez_shuffle을 기본적으로 사용합니다.

지원 중단되는 구성

다음 구성 속성은 HIVE-23354 이후 더 이상 사용되지 않으며 Amazon EMR 릴리스 6.11.0 이상에서는 더 이상 지원되지 않습니다.

명칭	기본값
`hive.mapred.reduce.tasks.speculative.execution`	`false`
`tez.am.speculation.enabled`	`false`

Amazon EMR 6.10.0 - Hive 릴리스 정보

Amazon EMR 6.10.0 - Hive 변경

유형	설명
Feature	IAM 패스스루(HiveCLI/Steps API)를 통해 Apache Hive 쿼리(쓰기)에 대한 AWS Lake Formation 기반 액세스 제어를 활성화합니다. HiveCLI
개선	로그 크기를 줄이려면 기본적으로 `hive.log.explain.output` 구성 비활성화
백포트	HIVE-26408: 벡터화: 스크래치 열의 할당 해제를 수정하고 하위 ConstantVectorExpression을 출력으로 재사용하지 않습니다.
백포트	HIVE-22269: HIVE-20703으로 인해 통계가 누락되어 동적 파티션을 포함하는 삽입 쿼리에서 잘못된 reducers 수 문제를 수정합니다.
백포트	HIVE-22891: LLAP 외 실행 모드에 대해 CombineHiveRecord에서 PartitionDesc 추출 건너뛰기
백포트	HIVE-23804: 이전 버전과의 호환성을 위해 Hive 메타스토어 스키마에 열 통계별 테이블에 대한 기본 데이터베이스 추가
백포트	HIVE-25277: 비용이 많이 드는 ListFile을 사용하는 클라우드 오브젝트 스토어에서 Hive 파티션 삭제 속도가 느림
백포트	HIVE-19202: HiveAggregate.isBucketedInput()dptj NullPointerException으로 인해 CBO 실패

유형	설명
백포트	HIVE-19048: beeline Initscript 오류(무시됨) 수정
백포트	HIVE-21085: 구체화된 뷰 레지스트리가 비외부 tez 세션 시작
백포트	HIVE-21675: 보기가 이미 있는 경우 CREATE VIEW IF NOT EXISTS에서 'OK'가 아닌 오류가 반환됩니다. 이는 Hive 2에서의 회귀입니다.
백포트	HIVE-21646: Tez: TezTasks가 스레드 로깅 컨텍스트를 이스케이프 처리하지 못하도록 방지
백포트	HIVE-22054: 디렉터리가 비어 있는지 확인하기 위한 재귀 나열 작업 방지
백포트	HIVE-16587: null 값이 중첩된 복합 유형을 삽입할 때 NPE
백포트	HIVE-22647: 기본적으로 세션 풀 활성화
백포트	HIVE-13288: DagUtils.localizeResource에서 혼란스러운 예외 메시지
백포트	HIVE-23870: WritableHiveCharObjectInspector.getPrimitiveJavaObject 및 HiveCharWritable에서 다중 텍스트 변환 최적화
백포트	HIVE-21498: Thrift를 0.13.0으로 업그레이드
백포트	HIVE-24378: 십진수 변환 전에 선행 및 후행 공백이 제거되지 않음
백포트	HIVE-21341: 합리적인 기본값: hive.server2.idle.operation.timeout 및 hive.server2.idle.session.timeout이 너무 높음

유형	설명
백포트	HIVE-22465: TezConfigurationFactory에서 ssl conf 추가
백포트	HIVE-24710: count(*)에 대해 CPU 및 IO 비용을 줄이도록 PTF 반복 최적화
백포트	HIVE-15406: 새로운 'trunc' 함수를 벡터화하는 방법 고려
백포트	HIVE-21541: HIVE-15406 버전에서 누락된 asf 헤더 수정
백포트	HIVE-24808: 구문 분석된 날짜 캐싱
백포트	HIVE-24746: PTF: 범위 계산 중 Timestamp ValueBoundaryScanner를 최적화할 수 있음
백포트	HIVE-25059: 복제 중에 변경 이벤트가 이름 바꾸기로 변환됨
백포트	HIVE-25142: 맵 조인 빠른 해시 테이블에서 재해시로 인해 대형 키가 손상됨
백포트	HIVE-23756: package.jdo 파일에 제약 조건 더 추가
백포트	HIVE-25150: HIVE-24378 일부로 수정된 공백 문자와 마찬가지로 십진수 변환 전에 탭 문자가 제거되지 않습니다.
백포트	HIVE-25093: date_format() UDF는 UTC 시간대로만 출력을 반환함
백포트	HIVE-25268: date_format udf는 현지 시간대가 UTC가 아닌 경우 1900년 이전 날짜에 대해 잘못된 결과 반환

유형	설명
백포트	HIVE-25338: 입력이 빈 경우 conv UDF의 AIOBE
백포트	HIVE-22400: UDF 분(시간 포함)에서 NULL 반환
백포트	HIVE-25058: PTF: 범위 계산 중 Timestamp ValueBoundaryScanner를 최적화할 수 있음, pt2 - isDistanceGreater
백포트	HIVE-25449: UTC 이외 시간대에서 tex 작업으로 실행 시 timezone datediff()에서 잘못된 출력 제공
백포트	HIVE-23688: 벡터화: IndexArrayOutOfBoundsException For 맵 유형 열에 null 값이 포함됨
백포트	HIVE-22247: 파티션 작업 출력이 빈 경우 HiveHFileOutputFormat에서 FileNotFoundException 발생
백포트	HIVE-25570: Hive는 명령 삽입 덮어쓰기 위치에 대한 승인을 위해 전체 URL 경로를 전송해야 함
백포트	HIVE-22903: 벡터화된 row_number()로 파티션 절에서 상수 표현식이 있는 경우 하나의 배치 이후 행 번호 재설정
백포트	HIVE-25549: PARTITION BY 또는 ORDER BY 절을 포함하는 표현식에서 기간 함수의 잘못된 결과
백포트	HIVE-25579: 덮어쓰기가 아닌 LOAD 덮어쓰기 추가

유형	설명
백포트	HIVE-25659: IN/(NOT IN)을 포함하는 메타스토어 직접 sql 쿼리는 SQL DB에서 허용하는 max 파라미터에 따라 분할되어야 함
백포트	HIVE-20502: 열 통계를 사용할 때 skewjoin_mapjoin10.q를 실행하는 중 발생하는 NPE 수정
백포트	HIVE-25765: skip.header.line.count 속성은 파일 크기가 더 크면 FetchOperator에서 각 블록의 행을 건너뜀
버그	특정 시나리오에서 `hive.stats.column.autogather` 및 `hive.groupby.skewindata` 모두 활성화된 경우 삽입 시 NPE 수정
버그	`mapred.tasktracker.expiry.interval` 값이 설정되지 않은 경우 NPE 수정

Amazon EMR 6.9.0 - Hive 릴리스 정보

Amazon EMR 6.9.0 - Hive 변경

유형	설명
업그레이드	Jetty를 9.4.48.v20220622로 업그레이드
업그레이드	Hadoop 3.3.3에 대한 지원
Feature	GCSC API를 사용하여 대화형 워크로드를 위해 Amazon EMR Hive와 Lake Formation 통합을 지원합니다.
Feature	Amazon EMR Hive와 Iceberg의 통합.

유형	설명
개선 사항	Amazon EMR 보안 구성을 사용하여 전송 중 암호화가 활성화된 경우 HiveServer2에서 SSL을 활성화합니다.
개선 사항	Hive EMRFS Amazon S3 최적화 커미터를 기본적으로 활성화합니다. 자세한 내용은 Hive EMRFS S3 최적화 커미터 활성화 섹션을 참조하세요.
개선 사항	SPARK-34210을 수정하기 위해 InputFormat의 mapred 버전만 상속하는 `HiveHBaseTableInputFormatV2` 를 추가합니다. `hive.hbase.inputformat.v2` 를 `true`로 설정하여 사용합니다.
개선 사항	TezAM을 종료하고 즉시 새로 실행하는 대신, hive.cli.tez.session.async를 사용하여 백그라운드에서 실행될 때까지 기다립니다. `hive.emr.cli.tez.session.open.timeout` 을 사용하여 이 제한 시간을 초 단위로 설정합니다.
개선 사항	hive.conf.restricted.list.append 옵션을 추가하여 기존의 제한된 구성 목록 `hive.conf.restricted.list` 에 쉼표로 구분된 구성을 추가합니다.
개선 사항	데이터베이스의 위치가 정의되지 않아 Hive 쿼리가 실패하는 경우 보다 명확한 오류 메시지를 제공합니다.
백포트	HIVE-24484: Hadoop을 3.3.1로, Tez를 0.10.2로 업그레이드
백포트	HIVE-22398: ShimLoader를 통해 YARN 큐 관리를 제거합니다.

유형	설명
백포트	HIVE-23190: LLAP: TezSpillRecord에 파일 시스템 객체를 전달하도록 IndexCache를 수정합니다.
백포트	HIVE-22185: HADOOP-15832로 인해 MiniYarn 클러스터를 사용한 테스트에서 문제가 발생합니다.
백포트	HIVE-21670: `mockito-all` 을 `mockito-core` 종속 항목으로 바꿉니다.
백포트	HIVE-24542: 업그레이드를 위해 Guava를 준비합ㄴ다.
백포트	HIVE-23751: QTest: HADOOP-16582 이후 정렬을 위해 `ProxyFileSystem` 에서 `#mkdirs()` 메서드를 재정의합니다.
백포트	HIVE-21603: Java 11 준비: powermock 버전을 업데이트합니다.
백포트	HIVE-24083: Hadoop 3.3.0에서 `hcatalog` 오류 발생: 인증 유형이 필요합니다.
백포트	HIVE-24282: 명시적으로 언급하지 않는 한 표시 열은 출력 열을 정렬해서는 안 됩니다.
백포트	HIVE-20656: 합리적인 기본값: 맵 어그리게이션 메모리 구성이 너무 공격적입니다.
백포트	HIVE-25443: 값이 1,024개가 넘는 경우 Arrow Serde는 복합 데이터 유형을 직렬화 및 직렬화 해제할 수 없음
백포트	HIVE-19792: orc를 1.5.2로 업그레이드하고 decimal_64 스키마 진화 테스트를 활성화합니다.

유형	설명
백포트	HIVE-20437: float, double, decimal에서 스키마 진화를 처리합니다.
백포트	HIVE-21987: Hive는 십진수로 주석이 달린 Parquet int32를 읽을 수 없습니다.
백포트	HIVE-20038: 버킷팅 및 파티셔닝되지 않은 테이블에서 업데이트 쿼리로 NPE가 발생합니다.

Amazon EMR 6.9.0 - Hive 알려진 문제

- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 HIVE-20703 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 HIVE-22269 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.

Amazon EMR 6.8.0 - Hive 릴리스 정보

Amazon EMR 6.8.0 - Hive 변경

유형	설명
개선 사항	msck 명령에서 파일 시스템 직접 호출을 줄입니다. 성능 향상(10,000개 이상의 파티션에서 최대 15~20배)
백포트	HIVE-20678: 호환성을 위해 HiveHBase TableOutputFormat에서 HiveOutputFormat을 구현해야 함
백포트	HIVE-21040: msck는 디렉터리 트리의 마지막 수준에서 불필요한 파일 목록을 표시함

유형	설명
백포트	HIVE-21460: 데이터를 로드한 후 select * 쿼리에서 잘못된 결과 생성
백포트	HIVE-21660: explode를 포함하는 union all 및 later view를 사용할 때 잘못된 결과
백포트	HIVE-22505: 잘못된 벡터화된 연산자 선택으로 ClassCastException 발생
백포트	HIVE-22513: 필터 작업에서 캐스팅된 열을 지속적으로 전파되면 잘못된 결과가 발생할 수 있음
백포트	HIVE-23435: 전체 외부 조인 결과에 행이 누락됨
백포트	HIVE-24209: 벡터화가 활성화된 경우 NOT BETWEEN 작업에 대한 잘못된 검색 인수 변환
백포트	HIVE-24934: VectorizedExpressions 주석이 GenericUDFSQCountCheck에 필요하지 않음
백포트	HIVE-25278: HiveProjectJoinTransposeRule에서 기간 함수에 대해 잘못된 변환을 수행할 수 있음
백포트	HIVE-25505: 헤더의 잘못된 결과. 첫 번째 줄이 빈 경우 skip.header.line.count
백포트	HIVE-26080: accumulo-core를 1.10.1로 업그레이드
백포트	HIVE-26235: 바이너리 열의 OR 조건이 빈 결과를 반환함
버그	시작 중에 stderr에서 여러 SLF4J 바인딩 경고 로그 수정

유형	설명
버그	파티션과 테이블이 다른 파일 시스템에 있을 때 잘못된 FS 오류로 SHOW TABLE EXTENDED 쿼리가 실패하는 문제를 수정합니다.

Amazon EMR 6.8.0 - Hive 알려진 문제

- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 HIVE-20703 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 HIVE-22269 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.

Amazon EMR 6.7.0 - Hive 릴리스 정보

Amazon EMR 6.7.0 - Hive 변경

유형	설명
Feature	Amazon EMR Hive와 LakeFormation 통합.
Feature	Hive EMRFS Amazon S3 최적화 커미터에 대한 추가 감사 로깅. Hive 구성: `hive.blobstore.output-committer.logging` , 기본값: `false`
Feature	삽입 시 삭제된 대상 디렉터리가 파티셔닝되지 않은 테이블 및 정적 파티션에 빈 select 결과를 덮어씁니다(Hive 2.x와 비슷한 동작). Hive 구성: `hive.emr.iow.clean.target.dir` , 기본값: `false`

유형	설명
버그	파티션 버킷 정렬과 함께 Hive EMRFS Amazon S3 최적화 커미터를 사용할 때 간헐적으로 발생하는 쿼리 오류를 수정했습니다.
업그레이드	Hive 버전을 3.1.3으로 업그레이드했습니다. 자세한 내용은 Apache Hive 3.1.3 릴리스 정보를 참조하세요.
업그레이드	Parquet을 1.12.2로 업그레이드했습니다.
백포트	HIVE-20065: 메타스토어는 jackson 1.x에 의존해서는 안 됨
백포트	HIVE-20071: jackson 2.x로 마이그레이션하고 사용 금지
백포트	HIVE-20607: TxnHandler는 직접 SQL 쿼리 실행을 위해 PreparedStatement를 사용해서는 안 됨
백포트	HIVE-20740: ObjectStore.setConf 메서드에서 글로벌 잠금 제거
백포트	HIVE-20961: NVL 구현 사용 중지
백포트	HIVE-22059: hive-exec jar에 (fasterxml) jackson 라이브러리가 포함되지 않음
백포트	HIVE-22351: TestObjectStore에서 잘못된 스레드 ObjectStore 사용 수정
백포트	HIVE-23534: 메시지 없이 MetaException 감지 시 RetryingMetaStoreClient#invoke에서 NPE 발생
백포트	HIVE-24048: jackson 구성 요소를 버전 2.10.latest - Hive으로 조화

유형	설명
백포트	HIVE-24768: 버전 교체를 위해 모든 곳에서 jackson-bom 사용
백포트	HIVE-24816: CVE-2020-25649 때문에 jackson 을 2.10.5.1 또는 2.11.0 이상으로 업그레이드
백포트	HIVE-25971: 캐시된 스레드 풀이 닫히지 않아 Tez 작업 종료가 지연됨
백포트	HIVE-26036: ObjectStore에서 getMTable()로 인해 NPE 발생

Amazon EMR 6.7.0 - Hive 알려진 문제

- 조인과 동일한 열에서 윈도우 함수를 사용하는 쿼리는 HIVE-25278에서 보고된 바와 같이 잘못된 변환으로 이어져 잘못된 결과를 생성하거나 쿼리에 실패할 수 있습니다. 해결 방법은 이러한 쿼리에 대해 쿼리 수준에서 CBO를 비활성화하는 것입니다. 수정 사항은 6.7.0 이후 Amazon EMR 릴리스에서 제공될 예정입니다. 자세한 내용은 AWS Support에 문의하세요.
- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 HIVE-20703 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 HIVE-22269 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.

Amazon EMR 6.6.0 - Hive 릴리스 정보

Amazon EMR 6.6.0 - Hive 변경

유형	설명
업그레이드	Parquet을 1.12.1로 업그레이드합니다.

유형	설명
업그레이드	jetty jars 버전을 9.4.43.v20210629로 업그레이드
버그	Hive 클러스터에서 LLAP가 활성화되었을 때 모든 태스크 및 코어 노드에 Hive가 설치되는 문제를 수정했습니다.
백포트	HIVE-25942: CVE-2021-29425 때문에 commons-io를 2.8.0으로 업그레이드
백포트	HIVE-25726: CVE-2020-13936 때문에 velocity를 2.3으로 업그레이드
백포트	HIVE-25680: HiveMetaStore 인증 모델을 사용하도록 #get_table_meta HiveMetaStore Server API에 권한을 부여합니다.
백포트	HIVE-25554: arrow 버전을 0.15로 업그레이드
백포트	HIVE-25242: vectorized.adaptor = chosen으로 쿼리 성능이 매우 저하됨
백포트	HIVE-25085: 메타스토어 클라이언트는 더 이상 세션 간에 공유되지 않음
백포트	HIVE-24827: Hive 집계 쿼리가 텍스트 파일이 아닌 잘못된 결과를 반환합니다.
백포트	HIVE-24683: 존재하지 않는 경로에 대해 Hadoop23Shim getFileId에서 NPE가 발생할 수 있음
백포트	HIVE-24656: 맵 및 배열 유형이 null인 쿼리에서 CBO 실패
백포트	HIVE-24556: 손자 항목이 없는 케이스에 대해 DefaultGraphWalker 최적화

유형	설명
백포트	HIVE-24408: Parquet을 1.11.1로 업그레이드
백포트	HIVE-24391: branch-3.1에서 FIX TestOrcFile 수정
백포트	HIVE-24362: 노드 수가 많은 트리에 대해 AST 트리 처리가 최적화된 상태가 아님
백포트	HIVE-24316: branch-3.1에서 ORC를 1.5.6에서 1.5.8로 업그레이드
백포트	HIVE-24307: property-file 및 -e 파라미터를 포함하는 Beeline 실패
백포트	HIVE-24245: 파티션에서 count 및 distinct를 포함하는 벡터화된 PTF로 인해 잘못된 결과가 생성됩니다.
백포트	HIVE-24224: 압축 파일에서 Tez 기반 Hive에 대한 머리글 및 바닥글을 건너뛰는 문제 수정
백포트	HIVE-24157: CAST 타임스탬프 ↔ 숫자에서 엄격한 모드 실패
백포트	HIVE-24113: GenericUDFToUnixTimeStamp에서 NPE
백포트	HIVE-23987: arrow 버전을 0.11.0으로 업그레이드
백포트	HIVE-23972: LLAP 외부 클라이언트에 외부 클라이언트 ID 추가
백포트	HIVE-23806: 스키마가 확장되는 경우 모든 파티션의 열 통계 상태 지우기 방지 이렇게 하면 alter table add columns 문의 런타임이 향상됩니다.

유형	설명
백포트	HIVE-23779: beeline 콘솔에서 BasicStatsTask 정보가 인쇄되지 않음
백포트	HIVE-23306: System.getProperty에 의해 설정된 구성이 있는 경우 RESET 명령이 작동하지 않음
백포트	HIVE-23164: 대몬(daemon) 스레드가 아니므로 서버가 올바르게 종료되지 않음
백포트	HIVE-22967: Tez 기반 Hive에 대해 hive.reloadable.aux.jars.path 지원
백포트	HIVE-22934: 오류 스트림에 대한 Hive 서버 대화형 로그 카운터
백포트	HIVE-22901: 변수 대체로 순환 참조에서 OOM이 발생할 수 있음
백포트	HIVE-22769: 압축된 텍스트 파일의 분할 생성 중 잘못된 쿼리 결과 및 쿼리 실패
백포트	HIVE-22716: ByteBuffer로의 읽기가 ParquetFooterInputFromCache에서 중단됨
백포트	HIVE-22648: Parquet을 1.11.0으로 업그레이드
백포트	HIVE-22640: Decimal64ColumnVector: 파티션 열 유형이 십진수인 경우 ClassCastException 발생
백포트	HIVE-22621: 불안정한 테스트 케이스: TestLlapSignerImpl.testSigning
백포트	HIVE-22533: 잠재적 LLAP 대몬(daemon) 웹 UI 취약성 수정

유형	설명
백포트	HIVE-22532: PTFPPD가 Rank, DenseRank 함수를 통해 제한을 잘못 푸시할 수 있음
백포트	HIVE-22514: HiveProtoLoggingHook에서 너무 많은 메모리를 소비할 수 있음
백포트	HIVE-22476: hive.fetch.task.conversio이 none으로 설정된 경우 Hive datediff 함수에서 일관되지 않은 결과를 제공함
백포트	HIVE-22429: hive 3에서 bucketing_version 1을 사용하는 마이그레이션된 클러스터형 테이블에서 삽입 시 bucketing_version 2 사용
백포트	HIVE-22412: explain 수행 시 StatsUtils에서 NPE 발생
백포트	HIVE-22360: 로드된 파일이 열이 테이블 스키마의 열보다 많은 경우 MultiDelimitserde가 마지막 열에 잘못된 결과를 반환함
백포트	HIVE-22332: Hive는 ORC-540 이후 유효한 스키마 진화 설정을 보장해야 함
백포트	HIVE-22331: 인수가 없는 unix_timestamp는 타임스탬프를 초 대신 밀리초 단위로 반환함
백포트	HIVE-22275: OperationManager.queryIdOperation이 여러 queryId를 올바르게 정리하지 않음
백포트	HIVE-22273: 임시 디렉터리가 제거되면 액세스 확인이 실패함
백포트	HIVE-22270: commons-io를 2.6으로 업그레이드

유형	설명
백포트	HIVE-22241: 내부 표현과 그레고리력-율리우스력 혼합 캘린더를 사용하여 날짜와 타임스탬프를 해석하도록 UDF 구현
백포트	HIVE-22241: 내부 표현과 그레고리력-율리우스력 혼합 항목을 사용하여 날짜와 타임스탬프를 해석하도록 UDF 구현
백포트	HIVE-22232: hive.order.columnalignment를 false로 설정한 경우 NPE
백포트	HIVE-22231: knox를 통한 큰 Hive쿼리에서 중단된 파이프 쓰기로 실패
백포트	HIVE-22221: LLAP 외부 클라이언트 - LlapBaseInputFormat#getSplits를 줄여야 함
백포트	HIVE-22208: 마스크 열이 있는 테이블에서 조인을 포함한 쿼리를 다시 작성할 때 예약된 키워드가 있는 열 이름의 이스케이프 처리가 해제됨
백포트	HIVE-22197: 일반 병합 조인에서 클래스 캐스트 예외가 발생합니다.
백포트	HIVE-22170: from_unixtime 및 unix_timestamp에서 사용자 세션 시간대를 사용해야 함
백포트	HIVE-22169: Tez: SplitGenerator에서 Tez에 없는 계획 파일을 찾으려고 함
백포트	HIVE-22168: llap cache hotpath에서 비용이 많이 드는 로깅 제거
백포트	HIVE-22161: UDF: org.apache.hadoop.hive.ql.udf.UDFType 클래스에서 FunctionRegistry 동기화

유형	설명
백포트	HIVE-22120: 특정 경계 조건에서 왼쪽 외부 맵 조인의 잘못된 결과 및 ArraOutOfBound 예외 수정
백포트	HIVE-22115: 속성이 false로 설정된 경우 쿼리 라우팅 어펜더 생성 방지
백포트	HIVE-22113: AMReporter 관련 RuntimeException에서 LLAP 종료 방지
백포트	HIVE-22106: 파티션 평가에 대한 교차 쿼리 동기화 제거
백포트	HIVE-22099: HIVE-20007 이후 여러 날짜 관련 UDF에서 율리우스력 날짜를 제대로 처리할 수 없음
백포트	HIVE-22037: OOM으로 인해 종료 시 HS2에서 로깅해야 함
백포트	HIVE-21976: Calcite HiveSortLimit에서 오프셋은 0 대신 null이어야 함
백포트	HIVE-21924: 머리글과 바닥글이 있어도 텍스트 파일 분할
백포트	HIVE-21913: GenericUDTFGetSplits는 LLAP와 동일한 방식으로 사용자 이름을 처리해야 함
백포트	HIVE-21905: FetchOperator 클래스 관련 제네릭 개선
백포트	HIVE-21902: HiveServer2 UI: jetty 응답 헤더에 X-Frame-Options가 필요함
백포트	HIVE-21888: hive.parquet.timestamp.skip.conversion 기본값을 true로 설정

유형	설명
백포트	HIVE-21868: CAST...FORMAT 벡터화
백포트	HIVE-21864: LlapBaseInputFormat#closeAll
백포트	HIVE-21863: WHEN 표현식을 위한 벡터라이저 유형 캐스팅 개선
백포트	HIVE-21862: ORC ppd가 타임스탬프와 함께 잘못된 결과를 생성함
백포트	HIVE-21846: LapDaemon 지표를 주기적으로 가져오는 스레드를 TeZam에 생성
백포트	HIVE-21837: 선택한 열의 값이 모두 null인 경우 MapJoin에서 예외 발생
백포트	HIVE-21834: 필터 조건을 단순화하기 위한 불필요한 직접 호출 방지
백포트	HIVE-21832: 평균 대기열, 지원, 응답 시간을 구하는 새로운 지표
백포트	HIVE-21827: SemanticAnalyzer에서 여러 직접 호출이 getTableObjectByName 메서드를 통과하지 않음
백포트	HIVE-21822: 새 API 메서드를 통해 LlapDaemon 지표 노출
백포트	HIVE-21818: CBO: TableRelOptHiveTable 복사 중 메타스토어 트래픽 포함
백포트	HIVE-21815: ORC 파일의 통계는 두 번 구문 분석됨
백포트	HIVE-21805: HiveServer2: 빠른 ShutdownHookManager API 사용

유형	설명
백포트	HIVE-21799: 조인 키가 집계 열에 있는 경우 DynamicPartitionPruningOptimization에서 NullPointerException 발생
백포트	HIVE-21794: 구체화된 뷰 파라미터를 sqlStdAuthSafeVarNameRegexes에 추가
백포트	HIVE-21768: JDBC: 포함되지 않은 UNION 쿼리의 기본 집합 접두사 제거
백포트	HIVE-21746: CBO가 비활성화된 상태로 동적으로 파티셔닝된 해시 조인 중 ArrayIndexOutOfBoundsException 발생
백포트	HIVE-21717: 이동 작업 중 디렉터리 이름 바꾸기에 실패했습니다.
백포트	HIVE-21685: IN 절이 여러 개 있는 쿼리의 단순화가 잘못됨
백포트	HIVE-21681: 여러 프라이머리 키에 대해 서식이 지정된 설명에 잘못된 정보가 표시됨
백포트	HIVE-21651: protobuf serde를 hive-exec로 이동합니다.
백포트	HIVE-21619: SQL explain extended에서 정밀도가 없는 타임스탬프 유형 인쇄
백포트	HIVE-21592: 표현식에 CONCAT가 포함된 경우 최적화된 SQL이 표시되지 않음
백포트	HIVE-21576: CAST...FORMAT 및 SQL:2016 날짜 시간 형식의 제한된 목록 도입

유형	설명
백포트	HIVE-21573: 인증이 delegationToken으로 설정된 경우 바이너리 전송은 보안 주체를 무시해야 함
백포트	HIVE-21550: TestObjectStore 테스트가 불안정함 - 요청된 시간 안에 잠금을 얻을 수 없음
백포트	HIVE-21544: 접는 동안 일정한 전파에서 coalesce, case, when 표현식 손상
백포트	HIVE-21539: 동일한 열에서 GroupBy + where 절 사용 시 잘못된 쿼리 재작성
백포트	HIVE-21538: Beeline: 콘솔 리더를 통해 암호 소스가 연결 파라미터에 전달되지 않음
백포트	HIVE-21509: LLAP가 손상된 열 벡터를 캐시하고 잘못된 쿼리 결과를 반환할 수 있음
백포트	HIVE-21499: AlreadyExistsException으로 인해 create 명령에 실패한 경우 레지스트리에서 함수를 제거해서는 안 됨
백포트	HIVE-21496: 정렬되지 않은 버퍼의 자동 크기 조정이 오버플로우될 수 있음
백포트	HIVE-21468: JDBC 스토리지 핸들러의 식별자 이름에서 대소문자 구분
백포트	HIVE-21467: 지원 중단된 junit.framework.Assert 가져오기 제거
백포트	HIVE-21435: LlapBaseInputFormat은 SubmitWorkRequestProto 빌드 중 TASK_ATTEMPT_ID conf(있는 경우)에서 작업 번호를 가져와야 함

유형	설명
백포트	HIVE-21389: HIVE-21247 이후 Hive 분배에서 javax.ws.rs-api.jar 누락
백포트	HIVE-21385: 분할할 수 없는 계산을 JDBC 소스로 푸시다운하는 기능을 비활성화할 수 있음
백포트	HIVE-21383: JDBC 스토리지 핸들러: 지정된 경우 카탈로그 및 스키마를 사용하여 테이블 검색
백포트	HIVE-21382: Group by 키 감소 최적화 - query23에서 키가 감소하지 않음
백포트	HIVE-21362: protobuf 파일에서 읽을 수 있도록 입력 형식 및 serde를 추가합니다.
백포트	HIVE-21340: CBO: SemiJoin에 입력되는 키가 아닌 열 정리
백포트	HIVE-21332: 잠긴 버퍼 대신 잠기지 않은 버퍼 제거
백포트	HIVE-21329: 운영자 파이프라인에 따라 사용자 지정 Tex 런타임에서 출력 버퍼 크기를 정렬하지 않음
백포트	HIVE-21295: StorageHandler는 Hive 규칙을 사용하여 날짜를 문자열로 변환해야 함
백포트	HIVE-21294: 벡터화: 1-reducer 셔플에서 객체 해시 함수를 건너뛸 수 있음
백포트	HIVE-21255: JdbcStorageHandler에서 QueryConditionBuilder 제거
백포트	HIVE-21253: JDBC StorageHandler에서 DB2 지원

유형	설명
백포트	HIVE-21232: LLAP: cache-miss 친화적 분할 선호도 공급자 추가
백포트	HIVE-21214: MoveTask : compareTempOrDuplicateFiles 파일 중복 제거에 대한 파일 크기 대신, attemptId 사용
백포트	HIVE-21184: 비용 정보와 함께 explain 및 explain 서식 지정된 CBO 계획 추가
백포트	HIVE-21182: 계획 중에 hive scratch dir 설정 건너뛰기
백포트	HIVE-21171: RPC가 켜져 있는 경우 tez용 스크래치 디렉터리 생성 건너뛰기
백포트	HIVE-21126: LlapBaseInputFormat#getSplit에서 세션 수준 쿼리 허용
백포트	HIVE-21107: 동적으로 분할된 해시 조인 중 'Cannot find field' 오류 발생
백포트	HIVE-21061: 빈 소스에 대해 IllegalStateException으로 CTAS 쿼리 실패
백포트	HIVE-21041: NPE, 논리적 계획에서 스키마를 가져오는 중 ParseException
백포트	HIVE-21013: JdbcStorageHandler가 Oracle에서 파티션 열을 찾지 못함
백포트	HIVE-21006: 재활용 기회가 있을 때 세미조인을 제거하도록 SharedWorkOptimizer 확장
백포트	HIVE-20992: hive.metastore.dbaccess.ssl.properties 구성을 더 유용한 구성으로 분할

유형	설명
백포트	HIVE-20989: JDBC - GetOperationStatus + 로그는 절전 모드를 통해 쿼리 진행을 차단할 수 있음
백포트	HIVE-20988: 프라이머리 키가 여러 열에 있는 group by 쿼리의 잘못된 결과
백포트	HIVE-20985: select 연산자 입력이 임시 열인 경우 벡터화로 일부가 출력으로 재사용될 수 있음
백포트	HIVE-20978: 'hive.jdbc.*'를 sqlStdAut hSafeVarNameRegexes에 추가해야 함
백포트	HIVE-20953: 생성 시 메타스토어에 추가할 수 없는 경우 함수 레지스트리에서 함수를 제거합니다.
백포트	HIVE-20952: VectorizationContext.java 정리
백포트	HIVE-20951: LLAP: 항상 Xms를 50%로 설정
백포트	HIVE-20949: 실제 계획에서 PKFK 카디널리티 추정 개선
백포트	HIVE-20944: 쿼리 컴파일 중에 통계를 검증하지 않음
백포트	HIVE-20940: Calcite의 유형 확인이 Hive보다 더 엄격한 케이스를 연결합니다.
백포트	HIVE-20937: Postgres jdbc 쿼리가 'LIMIT must not be negative'로 실패
백포트	HIVE-20926: bloom 필터 항목이 너무 높거나 통계가 없는 경우 세미조인 감소 힌트 실패
백포트	HIVE-20920: SQL 제약 조건을 사용하여 조인 재정렬 알고리즘 개선

유형	설명
백포트	HIVE-20918: Calcite에서 JDBC 연결로의 계산 푸시다운을 활성화하거나 비활성화하는 플래그
백포트	HIVE-20915: HoS 및 MR에서 동적 정렬 파티션 최적화를 사용할 수 있도록 설정
백포트	HIVE-20910: 동적 파티션 정렬 최적화로 인해 버킷팅된 테이블에서 삽입 실패
백포트	HIVE-20899: LLAP YARN 서비스의 키탭 URI는 HDFS만 지원하도록 제한됨
백포트	HIVE-20898: 시간 관련 함수의 경우 인수를 null을 허용하지 않는 유형으로 캐스팅할 수 없음
백포트	HIVE-20881: 지속적 전파는 투영을 단순화함
백포트	HIVE-20880: hive.stats.filter.in.min.ratio에 대한 기본값 업데이트
백포트	HIVE-20873: VectorHashKeyWrapperTwoLong에 대한 Murmur 해시를 사용하여 해시 충돌 감소
백포트	HIVE-20868: TezDummyOperator의 MapRecordProcessor 내 getFinalOp에 하위 연산자가 있는 경우 간헐적으로 SMB 조인 실패
백포트	HIVE-20853: llap 대몬(daemon) API에서 ShuffleHandler.registerDag 노출
백포트	HIVE-20850: 가능한 경우 프로젝션에서 차원 테이블로 케이스 조건 푸시
백포트	HIVE-20842: group by의 통계를 추정하기 위해 HIVE-20660에 도입된 로직 수정

유형	설명
백포트	HIVE-20839: 동적으로 분할된 해시 조인 중 'Cannot find field' 오류 발생
백포트	HIVE-20835: 제약 조건과 MV 재작성 간의 상호 작용으로 Calcite 플래너에서 루프가 생성될 수 있음
백포트	HIVE-20834: 캐시된 쿼리에서 SemanticAnalyzer에 대한 참조를 유지하는 Hive QueryResultCache 항목
백포트	HIVE-20830: 경우에 따라 JdbcStorageHandler 범위 쿼리 어설션 실패
백포트	HIVE-20829: JdbcStorageHandler 범위 분할에서 NPE 발생
백포트	HIVE-20827: 빈 배열에서 일관되지 않은 결과
백포트	HIVE-20826: 왼쪽의 조인 + group by를 왼쪽 세미 조인으로 변환하도록 HiveSemiJoin 규칙 개선
백포트	HIVE-20821: SUM0을 SUM + COALESCE 조합으로 재작성
백포트	HIVE-20815: JdbcRecordReader.next에서 예외가 발생하지 않아야 함
백포트	HIVE-20813: udf to_epoch_milli는 시간대 없이 타임스탬프를 지원해야 합니다.
백포트	HIVE-20804: 제약 조건이 있는 group by 최적화 추가 개선
백포트	HIVE-20792: 영역이 포함된 타임스탬프를 삽입하면 데이터가 잘림

유형	설명
백포트	HIVE-20788: 확장된 SJ 감소에서 필터 생성 시 열을 잘못 역추적할 수 있음
백포트	HIVE-20778: 계획의 모든 조인이 상관 해제 로직에 의해 생성된 경우 조인 재정렬이 트리거되지 않을 수 있음
백포트	HIVE-20772: LLAP에서 작업별 CPU 카운터 기록
백포트	HIVE-20768: Tumbling Window UDF 추가
백포트	HIVE-20767: 조인 연산자 간의 여러 프로젝트가 제약 조건을 사용한 조인 재정렬에 영향을 미칠 수 있음
백포트	HIVE-20762: NOTIFICATION_LOG 정리 간격이 60초로 하드코딩되어 너무 짧음
백포트	HIVE-20761: notification_sequence 테이블에서 업데이트에 대한 선택에 재시도 간격이 있으며, 재시도 수가 너무 작음
백포트	HIVE-20751: arrow 버전을 0.10.0으로 업그레이드
백포트	HIVE-20746: HiveProtoHookLogger가 하루 마감 시 파일을 닫지 않음
백포트	HIVE-20744: SQL 제약 조건을 사용하여 조인 재정렬 알고리즘 개선
백포트	HIVE-20740: ObjectStore.setConf 메서드에서 글로벌 잠금을 제거합니다. 이 cherrypick은 Hive 3.2 및 4.x에서 3.1.x에 적용 가능한 HIVE-20740을 백포트합니다.

유형	설명
백포트	HIVE-20734: Beeline: beeline-site.xml이 있고 hive CLI가 beeline으로 리디렉션되면 프롬프트를 표시하는 대신, 시스템 사용자 이름, 더미 암호를 사용해야 함
백포트	HIVE-20731: JdbcStorageHandler의 키스토어 파일이 승인되어야 함
백포트	HIVE-20720: JDBC 핸들러에 파티션 열 옵션 추가
백포트	HIVE-20719: hive.optimize.sort.dynamic.partition 최적화 및 벡터화가 켜져 있을 때 SELECT 문이 UPDATE 이후 실패함
백포트	HIVE-20718: 제약 조건이 있는 성능 CLI 드라이버 추가
백포트	HIVE-20716: hive.cbo.stats.correlated.multi.key.joins의 기본값을 true로 설정
백포트	HIVE-20712: HivePointLookupOptimizer에서 딥 케이스를 추출해야 함
백포트	HIVE-20710: 일정한 접기 작업으로 유형이 없는 null 상수를 생성할 수 없음
백포트	HIVE-20706: external_jdbc_table2.q 간헐적 실패
백포트	HIVE-20704: 다른 함수를 지원하도록 HivePreFilteringRule 확장
백포트	HIVE-20703: 동적 정렬 파티션 최적화를 비용 기반 의사 결정에 적용

유형	설명
백포트	HIVE-20702: 맵 조인 선택 중 데이터 구조 인식 추정으로 인한 오버헤드 고려
백포트	HIVE-20692: NOT x IS (NOT) [TRUE|FALSE] 표현식 접기 활성화
백포트	HIVE-20691: org.apache.hadoop.hive.cli. TestMiniLlapCliDriver.testCliDriver[cttl] 수정
백포트	HIVE-20682: 공유 sessionHive가 마스터 스레드에 의해 닫히면 비동기 쿼리 실행이 실패할 수 있음
백포트	HIVE-20676: HiveServer2: PrivilegeSynchronizer가 대몬(daemon) 상태로 설정되지 않음
백포트	HIVE-20660: 총 행 수를 소스 테이블로 바인딩하여 Group by 통계 추정을 개선할 수 있음
백포트	HIVE-20652: 서로 다른 두 데이터 소스를 jdbc 드라이버로 JdbcStorageHandler 푸시 조인
백포트	HIVE-20651: JdbcStorageHandler 암호를 암호화해야 함
백포트	HIVE-20649: Orc 라이터에 대한 LLAP 인식 메모리 관리자
백포트	HIVE-20648: LLAP: 벡터 group by 연산자에서 실행기당 메모리를 사용해야 함
백포트	HIVE-20646: IS NOT NULL인 경우 파티션 필터 조건이 메타스토어 쿼리로 푸시다운되지 않음
백포트	HIVE-20644: Hive 런타임 예외를 통해 민감한 정보 노출 방지
백포트	HIVE-20636: 외부 조인 후 null 값 수 추정 개선

유형	설명
백포트	HIVE-20632: 쿼리된 테이블에 구체화된 뷰를 생성하는 경우 get_splits UDF를 포함한 쿼리 실패
백포트	HIVE-20627: LockException으로 동시 비동기 쿼리가 간헐적으로 실패하고 메모리 누출이 발생함
백포트	HIVE-20623: 공유 작업: LLAP의 맵 조인 캐시 항목 공유 확장
백포트	HIVE-20619: 기본적으로 HiveServer2에 MultiDelimitSerDe 포함
백포트	HIVE-20618: 조인 선택 중 버킷팅되지 않은 테이블에 대해 BucketMapJoin이 선택될 수 있음
백포트	HIVE-20617: IN 표현식의 상수 유형을 올바른 유형으로 수정
백포트	HIVE-20612: CBO에 대한 새 조인 멀티키 상관관계 플래그 생성
백포트	HIVE-20603: 테이블 위치 파일 시스템을 변경한 후 파티션에 삽입할 때 'Wrong FS' 오류 발생
백포트	HIVE-20601: DbNotificationListener의 ALTER_PARTITION 이벤트에서 EnvironmentContext null
백포트	HIVE-20583: HiveConnection에서 kerberos 인증에만 정식 호스트 이름 사용
백포트	HIVE-20582: hive 프로토 로깅에서 hflush를 구성 가능하도록 설정

유형	설명
백포트	HIVE-20563: 벡터화: THEN/ELSE 유형 및 결과 유형이 다른 경우 CASE WHEN 표현식이 실패함
백포트	HIVE-20558: hive.hashtable.key.count.ad justment의 기본값을 0.99로 변경
백포트	HIVE-20552: LogicalPlan에서 보다 빠르게 스키마 가져오기
백포트	HIVE-20550: Hive 쿼리를 제출하기 위해 beeline을 사용하도록 WebHCat 전환
백포트	HIVE-20537: CBO와 Hive에서 상관되지 않은 다른 열을 포함하여 다중 열 조인 추정
백포트	HIVE-20524: 스키마 진화 검사가 Hive 버전 2에서 버전 3으로 이전 중단됨(ALTER TABLE VARCHAR에서 DECIMAL)
백포트	HIVE-20522: 필드의 null 허용 여부로 인해 HiveFilterSetOpTransposeRule에서 어설션 오류가 발생할 수 있음
백포트	HIVE-20521: HS2 doAs=true에는 hadoop.tm p.dir, MR 및 S3A 파일 시스템 관련 권한 문제가 있음
백포트	HIVE-20515: 결과 캐시 및 쿼리 임시 디렉터리를 사용할 때 쿼리 결과 비어 있음, 다른 파일 시스템의 결과 캐시 디렉터리
백포트	HIVE-20508: Hive는 'user@realm' 유형의 사용자 이름을 지원하지 않음
백포트	HIVE-20507: Beeline: beeline-site.xml의 모든 uris를 검색하도록 유틸리티 명령 추가

유형	설명
백포트	HIVE-20505: org.openjdk.jmh:jmh-core를 1.21로 업그레이드
백포트	HIVE-20503: 맵 조인 선택 시 데이터 구조 인식 추정 사용
백포트	HIVE-20498: 열 통계 자동 수집에 대해 날짜 유형 지원
백포트	HIVE-20496: 벡터화: 벡터화된 PTF IllegalStateException
백포트	HIVE-20494: HIVE-19440 이후 GenericUDFRestrictInformationSchema 손상됨
백포트	HIVE-20477: 표현식에 IN이 포함된 경우 최적화된 SQL이 표시되지 않음
백포트	HIVE-20467: 리소스 계획 및 삭제에서 IF NOT EXISTS/IF EXISTS 허용
백포트	HIVE-20462: 이미 보기가 있는 경우 'CREATE VIEW IF NOT EXISTS' 실패
백포트	HIVE-20455: security.authorization.PrivilegeSynchonizer.run에서 로그 생성
백포트	HIVE-20439: llap의 조인 선택 시 늘어난 메모리 제한 사용
백포트	HIVE-20433: 문자열에서 타임스탬프로의 암시적 변환 속도가 느림
백포트	HIVE-20432: 통계 추정을 위해 정수 유형에서 BETWEEN을 IN으로 재작성
백포트	HIVE-20423: NULLS LAST를 기본 null 정렬로 설정

유형	설명
백포트	HIVE-20418: LLAP IO는 열을 선택하지 않은 쿼리에 대해 행 인덱스가 올바르게 비활성화된 ORC 파일을 처리하지 못할 수 있음
백포트	HIVE-20412: HiveMetaHook의 NPE
백포트	HIVE-20406: 중첩된 Coalesce로 인해 잘못된 결과 생성
백포트	HIVE-20399: 완전히 검증되지 않은 사용자 지정 테이블 위치를 포함하는 CTAS는 MM 테이블에서 실패
백포트	HIVE-20393: 세미조인 감소: markSemiJoinForDPP의 일관되지 않은 동작
백포트	HIVE-20391: 집계 함수를 분해할 때 HiveAggregateReduceFunctionsRule에서 잘못된 반환 유형을 추론할 수 있음
백포트	HIVE-20383: hive 프로토 이벤트 후크에서 잘못된 대기열 이름 및 동기화 문제.
백포트	HIVE-20367: 벡터화: PTF AVG, MAX, MIN, SUM에 대한 스트리밍 지원
백포트	HIVE-20366: TPC-DS query78 통계 추정이 꺼짐, null filte
백포트	HIVE-20364: hive.map.aggr.hash.min.reduction의 기본값 업데이트
백포트	HIVE-20352: 벡터화: 그룹화 기능 지원
백포트	HIVE-20347: hive.optimize.sort.dynamic.partition이 파티셔닝된 CTAS 및 MV에서 작동해야 함

유형	설명
백포트	HIVE-20345: 다른 직접 호출에서 테이블이 삭제되면 데이터베이스 삭제가 중단될 수 있음
백포트	HIVE-20343: Hive 3: CTAS는 transactional_properties를 인식하지 않음
백포트	HIVE-20340: timestamp 함수의 출력을 문자열로 사용할 때 Druid 에는 타임스탬프에서 문자열로의 명시적 CAST가 필요함
백포트	HIVE-20339: 벡터화: 불필요한 제한을 높여 RANK가 있는 일부 PTF가 벡터화되지 않음
백포트	HIVE-20337: CachedStore: getPartitionsByExpr이 파티션 목록을 잘못 채움
백포트	HIVE-20336: 구체화된 뷰에 대한 마스킹 및 필터링 정책
백포트	HIVE-20326: NO RELY 대신 기본값으로 RELY를 포함하는 제약 조건 생성
백포트	HIVE-20321: 벡터화: 1 col VectorHashKeyWrapper의 메모리 크기를 <1 CacheLine로 감소
백포트	HIVE-20320: hive.optimize.remove.sq_count_check 플래그 켜기
백포트	HIVE-20315: 벡터화: 더 많은 NULL 및 잘못된 결과 문제 수정, 불필요한 캐스트 및 변환 방지
백포트	HIVE-20314: 구체화된 뷰 재작성에 파티션 정리 포함

유형	설명
백포트	HIVE-20312: arrow 클라이언트에서 LlapOutputFormatService와 함께 자체 BufferAllocator를 사용하도록 허용
백포트	HIVE-20302: LLAP: IO에서 벡터화되지 않은 실행이 가상 열 무시(ROW__ID 포함)
백포트	HIVE-20300: VectorFileSinkArrowOperator
백포트	HIVE-20299: LLAP 서명자 유닛 테스트에서 잠재적 경합
백포트	HIVE-20296: 더 정교한 컨텍스트에서 추출할 수 있도록 HivePointLookupOptimizerRule 개선
백포트	HIVE-20294: 벡터화: COALESCE 및 ELT에서 NULL 및 잘못된 결과 문제 수정
백포트	HIVE-20292: 기본 제약 조건이 정의된 tpcds query93의 조인 정렬이 잘못됨
백포트	HIVE-20290: GetSplits 중에 버퍼를 할당하지 않도록 ArrowColumnarBatchSerDe 초기화 지연
백포트	HIVE-20281: SharedWorkOptimizer가 'operator cache contents and actual plan differ'로 실패
백포트	HIVE-20277: 벡터화: BOOLEAN을 반환하는 케이스 표현식은 FILTER에서 지원되지 않음
백포트	HIVE-20267: 로그 수준을 동적으로 구성하는 양식을 포함하도록 WebUI 확장
백포트	HIVE-20263: HiveReduceExpressionsWithStatsRule 변수의 오타
백포트	HIVE-20260: 다른 열의 필터로 행 수가 변경될 때 열의 NDV를 조정해서는 안 됨

유형	설명
백포트	HIVE-20252: 세미조인 감소: 작은 테이블에 맵 조인 업스트림이 있는 경우 세미 조인 브랜치로 인한 주기가 감지되지 않을 수 있습니다.
백포트	HIVE-20245: 벡터화: BETWEEN 및 IN에서 NULL 및 잘못된 결과 문제 수정
백포트	HIVE-20241: CTAS 문의 파티셔닝 사양 지원
백포트	HIVE-20240: 세미조인 감소: 로컬 변수를 사용하여 외부 테이블 조건 검사
백포트	HIVE-20226: 요청 maxEvents가 테이블의 max_rows를 초과하므로 HMS getNextNotification에서 예외 발생
백포트	HIVE-20225: 테라데이타 바이너리 형식을 지원하는 Serde
백포트	HIVE-20213: Calcite를 1.17.0으로 업그레이드
백포트	HIVE-20212: http 모드의 Hiveserver2가 default.General.open_connections 지표를 잘못 생성함
백포트	HIVE-20210: 파티션이 아닌 열에 대한 필터 및 변환이 최소일 때 단순 가져오기 최적화 도구에서 MapReduce로 이어져야 함
백포트	HIVE-20209: repl 덤프에서 메타스토어 연결의 첫 번째 시도 실패
백포트	HIVE-20207: 벡터화: 필터 및 비교에서 NULL 및 잘못된 결과 문제 수정
백포트	HIVE-20204: IN 작업 중 유형 변환

유형	설명
백포트	HIVE-20203: Arrow SerDe에서 DirectByteBuffer 누출
백포트	HIVE-20197: 벡터화: DECIMAL_64 테스트 추가, 날짜, 산술, 타임스탬프 산술 추가, GROUP BY 집계 더 추가
백포트	HIVE-20193: explain plan json에 cboInfo가 없음
백포트	HIVE-20192: 메타스토어가 포함된 HS2에서 JDOPersistenceManager 객체 유출
백포트	HIVE-20183: 소스 테이블에 빈 버킷이 있는 경우 버킷팅된 테이블에서 삽입하면 데이터가 손실될 수 있음
백포트	HIVE-20177: 벡터화: GroupBy 스트리밍 모드에서 KeyWrapper 할당 감소
백포트	HIVE-20174: 벡터화: GROUP BY 집계 함수에서 NULL 및 잘못된 결과 문제 수정
백포트	HIVE-20172: 원격 메타스토어에 연결하려는 중 StatsUpdater에서 GSS 예외로 실패
백포트	HIVE-20153: Count 및 Sum UDF가 Hive 2 이상에서 더 많은 메모리를 소비함
백포트	HIVE-20152: repl 덤프가 실패하면 db 상태를 재설정하여 테이블 이름을 바꿀 수 있음
백포트	HIVE-20149: TestHiveCli 실패 또는 제한 시간 초과
백포트	HIVE-20130: 정보 스키마 동기화 도구에 대한 향상된 로깅

유형	설명
백포트	HIVE-20129: orc 테이블의 위치 기반 스키마 진화로 되돌리기
백포트	HIVE-20118: SessionStateUserAuthenticator.getGroupNames
백포트	HIVE-20116: TezTask에서 상위 로거 사용 중
백포트	HIVE-20115: ACID 테이블은 분석을 위해 바닥글 스캔을 사용해서는 안 됨
백포트	HIVE-20103: WM: 하나 이상이 사용된 경우 DAG 카운터만 집계
백포트	HIVE-20101: BloomKFilter: 로컬 byte[] 배열 완전히 사용 금지
백포트	HIVE-20100: OpTraits : 불일치 감지 시 Select Optraits를 중지해야 함
백포트	HIVE-20098: 통계: 날짜 열 파티션 통계를 가져오는 중 NPE
백포트	HIVE-20095: jdbc 외부 테이블로 계산을 푸시하는 기능 수정
백포트	HIVE-20093: LlapOutputFomatService: 회계에 대해 Netty와 함께 ArrowBuf 사용
백포트	HIVE-20090: 새로운 기회를 발견할 수 있도록 세미조인 감소 필터의 생성 확장
백포트	HIVE-20088: Beeline 구성 위치 경로가 잘못 구성됨
백포트	HIVE-20082: HiveDecimal에서 문자열로 변환해도 십진수의 형식이 올바르게 지정되지 않음

유형	설명
백포트	HIVE-20069: DPP 및 Semijoin 최적화의 경우 재최적화 수정
백포트	HIVE-20051: 임시 테이블에 대한 인증 건너뛰기
백포트	HIVE-20044: Arrow Serde는 문자 값을 채우고 빈 문자열을 올바르게 처리해야 함
백포트	HIVE-20028: 메타스토어 클라이언트 캐시 구성이 잘못 사용됨
백포트	HIVE-20025: HiveProtoLoggingHook에서 생성한 이벤트 파일 정리
백포트	HIVE-20020: Hive contrib jar이 라이브러리에 없어야 함
백포트	HIVE-20013: to_date 함수의 날짜 유형에 암시적 캐스트 추가
백포트	HIVE-20011: proto 로깅 후크에서 추가 모드 해제
백포트	HIVE-20005: acid_table_stats, acid_no_buckets 등 - 브랜치에서 쿼리 결과 변경
백포트	HIVE-20004: ConvertDecimal64ToDecimal에 사용된 잘못된 규모로 잘못된 결과 생성
백포트	HIVE-19995: ACID 테이블의 행 트래픽 집계
백포트	HIVE-19993: 열 이름으로도 나타나는 테이블 별칭을 사용할 수 없음

유형	설명
백포트	HIVE-19992: 벡터화: HIVE-19951 후속 조치 --> 데이터 유형 변환이 암시적이지 않은 경우에만 ORC에 대한 인코딩된 LLAP I/O를 비활성화하도록 SchemaEvolution.isOnlyImplicitConversion에 직접 호출 추가
백포트	HIVE-19989: 메타스토어가 HADOOP2 지표에 잘못된 애플리케이션 이름을 사용함
백포트	HIVE-19981: 테이블 삭제 시 데이터를 삭제하도록 HiveStrictManagedMigration 유틸리티에서 외부 테이블로 변환된 관리형 테이블을 설정해야 함
백포트	HIVE-19967: SMB 조인: PTFOperator ala GBY Op에 대한 Optraits 필요
백포트	HIVE-19935: Hive WM 세션 강제 종료: LLAP 작업 수를 업데이트하지 못함
백포트	HIVE-19924: Repl 로드에 의해 실행되는 태그 distcp 작업
백포트	HIVE-19891: 사용자 지정 파티션 디렉터리가 있는 외부 테이블에 삽입하면 데이터가 손실될 수 있음
백포트	HIVE-19850: Tez의 동적 파티션 정리로 인해 'No work found for tablescan' 오류 발생
백포트	HIVE-19806: 테스트 결과에서 불안정 문제를 방지하기 위해 qtests 출력 정렬
백포트	HIVE-19770: select에 동일한 열이 여러 개 있는 쿼리에 대한 CBO 지원

유형	설명
백포트	HIVE-19769: DB 및 테이블 이름에 대한 전용 객체 생성
백포트	HIVE-19765: BlobstoreCliDriver에 Parquet 특정 테스트 추가
백포트	HIVE-19759: 불안정 테스트: TestRpc#testServerPort
백포트	HIVE-19711: Hive 스키마 도구 리팩터링
백포트	HIVE-19701: getDelegationTokenFromMetaStore는 동기화하지 않아도 됨
백포트	HIVE-19694: 구체화된 뷰 생성 문은 MV의 SQL 문을 실행하기 전에 MV 이름 충돌을 확인해야 합니다.
백포트	HIVE-19674: Group by 십진수 상수는 Druid 테이블로 푸시다운
백포트	HIVE-19668: 힙의 30% 넘게 중복 org.antlr.runtime.CommonToken 및 중복 문자열에서 소비
백포트	HIVE-19663: LLAP IO 보고서 생성 리팩터링
백포트	HIVE-19661: Hive UDF를 Re2J 정규식 엔진을 사용하도록 전환
백포트	HIVE-19628: LLAP testSigning에서 잠재적 NPE
백포트	HIVE-19568: 액티브/패시브 HS2 HA: 패시브 HS2 인스턴스에 대한 직접 연결을 허용하지 않음

유형	설명
백포트	HIVE-19564: 벡터화: 산술의 NULL 및 잘못된 결과 문제 수정
백포트	HIVE-19552: TestMiniDruidKafkaCliDriver #druidkafkamini_basic.q 활성화
백포트	HIVE-19432: hive에 데이터베이스 및 테이블이 너무 많은 경우 GetTablesOperation 속도가 느려짐
백포트	HIVE-19360: CBO: QueryPlan 객체에 'optimizedSQL' 추가
백포트	HIVE-19326: 통계 자동 수집: UNION 쿼리 중 잘못된 집계
백포트	HIVE-19313: TestJdbcWithDBTokenStoreNoDoAs 테스트 실패
백포트	HIVE-19285: MetaDataOperation의 하위 클래스에 로그 추가
백포트	HIVE-19235: Minimr 테스트를 위한 골든 파일 업데이트
백포트	HIVE-19104: 재시도를 통해 메타스토어 테스트를 시작하는 경우 인스턴스가 독립적이어야 함
백포트	HIVE-18986: 테이블에 열이 많은 경우 테이블 이름을 바꿀 때 dataNucleus에서 java.lang.StackOverflowError 발생
백포트	HIVE-18920: CBO: 첫 번째 쿼리에 Janino 공급자 초기화

유형	설명
백포트	HIVE-18873: HiveInputFormat에서 MR에 대한 조건자 푸시다운을 자동으로 건너뛰면 스토리지 핸들러에서 잘못된 결과를 생성할 수 있음
백포트	HIVE-18871: hive.aux.jars.path를 hdfs://로 설정했기 때문에 tez 기반 hive에서 실행 오류 발생
백포트	HIVE-18725: 잘못된 열 참조가 있는 경우 하위 쿼리에 대한 오류 처리 개선
백포트	HIVE-18696: 다음과 같은 경우 HiveMetaStore.add_partitions_Core 메서드에서 파티션 폴더가 제대로 정리되지 않을 수 있음
백포트	HIVE-18453: ACID: 'CREATE TRANSACTIONAL TABLE' 구문을 추가하여 ACID ORC 및 Parquet 지원 통합
백포트	HIVE-18201: sq_count_chec에 대해 XPROD_EDGE 비활성화
백포트	HIVE-18140: 기본 통계 혼합 케이스에서 분할된 테이블 통계가 잘못될 수 있음
백포트	HIVE-17921: LLAP에서 구문을 사용한 집계에서 잘못된 결과 생성
백포트	HIVE-17896: TopNKey: 벡터화 가능한 독립 실행형 TopnKey 연산자 생성
백포트	HIVE-17840: transactionalListeners.notifyEvent에 실패한 경우 HiveMetaStore에서 예외 발생
백포트	HIVE-17043: 나중에 참조하지 않을 경우 group by 키에서 고유하지 않은 열 제거

유형	설명
백포트	HIVE-17040: FK 관계가 있는 상태에서 조인 제거
백포트	HIVE-16839: 동일한 파티션을 동시에 변경하는 경우 openTransaction/commitTransaction에 대한 균형이 맞지 않는 직접 호출
백포트	HIVE-16100: 동적 정렬 파티션 최적화 도구에서 형제 연산자 손실
백포트	HIVE-15956: 파티션의 많은 삭제 시 StackOverflowError 발생
백포트	HIVE-15177: kerberos 인증 유형이 fromSubject로 설정되고 보안 주체가 _HOST를 포함하는 경우 hive에서 인증 실패
백포트	HIVE-14898: HS2는 빈 인증 헤더 오류에 대한 콜 스택을 로깅하지 않아야 함
백포트	HIVE-14493: 구체화된 뷰에 대한 파티셔닝 지원
백포트	HIVE-14431: COALESCE를 CASE로 인식
백포트	HIVE-13457: 모니터링 정보를 위한 HS2 REST API 엔드포인트 생성
백포트	HIVE-12342: hive.optimize.index.filter의 기본값을 true로 설정
백포트	HIVE-10296: Hive가 메타스토어에서 다중 조인 쿼리를 실행할 때 캐스트 예외가 관찰됨
백포트	HIVE-6980: direct sql을 사용하여 테이블 삭제

Amazon EMR 6.6.0 - Hive 구성 변경

- OSS 변경 HIVE-20703의 일환으로 동적 파티션을 정렬하는 속성 (`hive.optimize.sort.dynamic.partition`)이 `hive.optimize.sort.dynamic.partition.threshold`로 바뀌었습니다.

  `hive.optimize.sort.dynamic.partition.threshold` 구성에는 다음과 같은 잠재적 값이 있습니다.

값	설명
`0` (기본값)	ORC 파일을 사용할 때 동적 파티션을 정렬하기 위한 최적화를 비용 기반으로 결정합니다. INSERT 쿼리에 허용되는 최대 라이터 수는 (실행기/컨테이너 메모리) * (orc가 차지하는 메모리 비율) 값을 단일 라이터가 사용한 최대 메모리(스트라이프 크기)로 나눈 값을 기준으로 계산됩니다.
`-1`	동적 파티션을 완전히 정렬하기 위해 최적화를 비활성화합니다.
1	동적 파티션의 글로벌 정렬을 활성화합니다. 이렇게 하면 reducer의 각 파티션 값에 대해 레코드 라이터를 하나만 열어 두므로 reducer에 가해지는 메모리 부담을 줄일 수 있습니다.
2 (또는 더 큰 정수)	지정된 정수를 최대 라이터 수의 임계값으로 사용하도록 Hive에 지시합니다.

Amazon EMR 6.6.0 - Hive 알려진 문제

- 조인과 동일한 열에서 윈도우 함수를 사용하는 쿼리는 HIVE-25278에서 보고된 바와 같이 잘못된 변환으로 이어져 잘못된 결과를 생성하거나 쿼리에 실패할 수 있습니다. 해결 방법으로 이러한 쿼리에 대해 쿼리 수준에서 CBO를 비활성화할 수 있습니다. 자세한 내용은 AWS Support에 문의하세요.
- Amazon EMR 6.6.0에는 Hive 소프트웨어 버전 3.1.2가 포함되어 있습니다. Hive 3.1.2에서는 머리글과 바닥글을 포함하는 경우 텍스트 파일을 분할하는 기능을 도입했습니다(HIVE-21924). Apache Tez App Master는 각 파일을 읽고 데이터 범위에서 오프셋 지점을 결정합니다. 쿼리에서 작은 텍스트 파일을 많이 읽을 경우 이러한 동작이 결합되어 성능이 저하될 수 있습니다. 이 문제를 해결하려면 `CombineHiveInputFormat`을 사용하고 다음 속성을 구성하여 최대 분할 크기를 조정합니다.

```
SET hive.tez.input.format=org.apache.hadoop.hive.ql.io.CombineHiveInputFormat;
```

```
SET mapreduce.input.fileinputformat.split.maxsize=16777216;
```

- Amazon EMR 6.6.0~6.9.x에서는 ORDER BY 또는 SORT BY 절과 동적 파티션에서의 INSERT 쿼리에 항상 두 개의 reducer가 있습니다. 이 문제는 OSS 변경 HIVE-20703 때문에 발생하며, 이로 인해 비용 기반 의사 결정에 따라 동적 정렬 파티션 최적화가 적용됩니다. 워크로드에 동적 파티션 정렬이 필요하지 않은 경우 새 기능을 비활성화하고 올바르게 계산된 reducer 수를 가져오도록 `hive.optimize.sort.dynamic.partition.threshold` 속성을 `-1`로 설정하는 것이 좋습니다. 이 문제는 OSS Hive에서 HIVE-22269 일부로 수정되었으며 Amazon EMR 6.10.0에서 수정되었습니다.

Hudi

Apache Hudi는 증분 데이터 처리 및 데이터 파이프라인 개발을 간소화하는 데 사용되는 오픈 소스 데이터 관리 프레임워크로, 레코드 수준의 삽입, 업데이트, 업서트 및 삭제 기능을 제공합니다. Upsert란 레코드가 존재하지 않은 경우 기존 데이터 세트에 레코드를 삽입하거나 레코드가 존재하는 경우 레코드를 업데이트할 수 있는 기능입니다. Hudi를 이용해 Amazon S3에 데이터를 배치하는 방법을 효율적으로 관리함으로써 거의 실시간으로 데이터를 수집하고 업데이트할 수 있습니다. Hudi에서는 데이터 세트에 수행된 작업의 메타데이터가 세심히 관리되므로 작업의 원자성 및 일관성을 보장하는 데 도움이 됩니다.

Hudi는 Apache Spark, Apache Hive, Presto와 통합됩니다. Amazon EMR 릴리스 버전 6.1.0 이상에서 Hudi는 Trino(PrestoSQL)와도 통합되었습니다.

Amazon EMR 릴리스 버전 5.28.0 이상에서 EMR은 Spark, Hive, Presto 또는 Flink가 설치된 경우 기본적으로 Hudi 구성 요소를 설치합니다. Spark 또는 Hudi DeltaStreamer 유틸리티를 사용하여 Hudi 데이터 세트를 만들거나 업데이트할 수 있습니다. Hive, Spark, Presto 또는 Flink를 사용하여 Hudi 데이터 세트를 대화식으로 쿼리하거나 증분 풀을 사용하여 데이터 처리 파이프라인을 구축할 수 있습니다. 증분 풀이란 두 작업 간에 변경된 데이터만 가져올 수 있는 기능입니다.

이러한 기능을 통해 다음 사용 사례에서 Hudi를 사용하여 유리합니다.

- 특정 데이터 삽입 및 업데이트 이벤트가 필요한 센서 및 기타 사물 인터넷(IoT) 디바이스에서의 스트리밍 데이터 작업.
- 사용자가 데이터 사용 방법에 대한 동의를 잊어버릴 수 있거나 수정할 수도 있는 애플리케이션의 데이터 프라이버시 규정 준수.
- 시간이 지남에 따라 데이터 세트에 변경 사항을 적용할 수 있는 데이터 캡처 변경(CDC) 시스템 구현.

다음 테이블에는 Amazon EMR이 Hudi를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Hudi의 버전이 나열되어 있습니다.

이 릴리스에서 Hudi와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Hudi 버전 정보

Amazon EMR 릴리스 레이블	Hudi 버전	Hudi와 함께 설치된 구성 요소
emr-7.9.0	Hudi 0.15.0-amzn-6	Not available.

다음 테이블에는 Amazon EMR이 Hudi를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Hudi의 버전이 나열되어 있습니다.

이 릴리스에서 Hudi와 함께 설치된 구성 요소의 버전은 [릴리스 6.15.0 구성 요소 버전]을 참조하세요.

emr-6.15.0용 Hudi 버전 정보

Amazon EMR 릴리스 레이블	Hudi 버전	Hudi와 함께 설치된 구성 요소
emr-6.15.0	Hudi 0.14.0-amzn-0	Not available.

> **Note**
>
> Amazon EMR 릴리스 6.8.0은 [Apache Hudi] 0.11.1과 함께 제공되지만, Amazon EMR 6.8.0 클러스터는 Hudi 0.12.0의 오픈 소스 `hudi-spark3.3-bundle_2.12`와도 호환됩니다.

다음 테이블에는 Amazon EMR이 Hudi를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Hudi의 버전이 나열되어 있습니다.

이 릴리스에서 Hudi와 함께 설치된 구성 요소의 버전은 [릴리스 5.36.2 구성 요소 버전]을 참조하세요.

emr-5.36.2용 Hudi 버전 정보

Amazon EMR 릴리스 레이블	Hudi 버전	Hudi와 함께 설치된 구성 요소
emr-5.36.2	Hudi 0.10.1-amzn-1	Not available.

주제

- Hudi 작동 방식
- Amazon EMR 기반 Hudi 사용 시 고려 사항 및 제한 사항
- Hudi가 설치된 클러스터 생성
- Hudi 데이터 세트 작업
- Hudi CLI 사용
- Hudi 릴리스 기록

Hudi 작동 방식

Amazon EMR에서 Hudi를 사용하는 경우 Spark 데이터 소스 API 또는 Hudi DeltaStremer 유틸리티를 사용하여 데이터 세트에 데이터를 쓸 수 있습니다. Hudi는 데이터 세트를 기존 Hive 테이블과 유사한 `basepath` 아래 파티셔닝된 디렉터리 구조로 구성합니다. 데이터가 이러한 디렉터리의 파일로 배치되는 방법에 대한 세부 사항은 선택한 데이터 세트 유형에 따라 다릅니다. CoW(쓸 때 복사) 또는 MoR(읽을 때 병합) 중 하나를 선택할 수 있습니다.

데이터 세트 유형에 상관없이 데이터 세트의 각 파티션은 `basepath`와(과) 관련된 `partitionpath`에 의해 고유하게 식별됩니다. 각 파티션 내에서 레코드는 여러 데이터 파일로 배포됩니다. 자세한 내용은 Apache Hudi 설명서의 파일 관리를 참조하십시오.

Hudi의 각 작업에는 해당하는 커밋이 있으며, 이 커밋은 일정하게 증가하는 타임스탬프(인스턴트라고도 함)로 식별됩니다. Hudi는 데이터 세트에서 수행된 일련의 모든 작업을 타임라인으로 유지합니다. Hudi는 타임라인을 통해 리더와 라이터 간 스냅샷 격리를 제공하고 이전 시점으로 롤백할 수 있도록 합니다. Hudi 레코드 및 작업 상태에 대한 자세한 내용은 Apache Hudi 설명서에서 Timeline을 참조하세요.

데이터 세트 스토리지 유형 이해: 쓸 때 복사 및 읽을 때 병합

Hudi 데이터 세트를 만들 때 데이터 세트를 쓸 때 복사 또는 읽을 때 병합으로 지정합니다.

- 쓸 때 복사(CoW) - 데이터가 열 기반 형식(Parquet)으로 저장되며, 각 업데이트마다 쓰기 중에 새 버전의 파일을 만듭니다. CoW가 기본 스토리지 유형입니다.
- 읽을 때 병합(MoR) - 데이터가 열 기반 형식(Parquet)과 행 기반(Avro) 형식을 조합하여 저장됩니다. 업데이트는 행 기반 delta 파일에 기록되며 새 버전의 열 형식 파일을 작성할 때 필요에 따라 압축됩니다.

CoW 데이터 세트를 사용하면 레코드에 대한 업데이트가 있을 때마다 레코드가 포함된 파일이 업데이트된 값으로 다시 작성됩니다. MoR 데이터 세트를 사용하면 Hudi는 업데이트가 있을 때마다 변경된 레코드에 대한 행만 씁니다. MoR은 읽기 수행이 적고 쓰기 또는 변경이 많은 워크로드에 더 적합합니다. CoW는 자주 변경되지 않는 데이터에서 읽기 수행이 많은 워크로드에 더 적합합니다.

Hudi는 데이터에 액세스하기 위한 세 가지 논리 뷰를 제공합니다:

- 읽기 최적화 보기 - CoW 테이블의 최신 커밋된 데이터 세트와 MoR 테이블의 최신 압축 데이터 세트를 제공합니다.

- 증분 보기 - 다운스트림 작업을 내보내고 워크플로를 추출, 변환, 로드(ETL)하기 위해 CoW 데이터 세트의 두 작업 간에 변경 스트림을 제공합니다.
- 실시간 보기 - 열 및 행 기반 파일을 인라인으로 병합하여 MoR 테이블에서 커밋된 최신 데이터를 제공합니다.

읽기 최적화 뷰를 쿼리하면 쿼리는 압축된 모든 데이터를 반환하지만 최신 델타 커밋은 포함하지 않습니다. 이러한 데이터를 쿼리하면 읽기 성능은 뛰어나지만 최신 데이터는 생략됩니다. 실시간 뷰를 쿼리하는 경우 Hudi는 읽을 때 압축된 데이터를 델타 커밋과 병합합니다. 최신 데이터는 쿼리할 수 있지만 병합의 컴퓨팅 오버헤드로 인해 쿼리 성능은 떨어집니다. 압축된 데이터 또는 실시간 데이터를 쿼리할 수 있으므로 쿼리할 때 성능과 유연성 중에서 선택할 수 있습니다.

스토리지 유형 간의 장단점에 대한 자세한 내용은 Apache Hudi 설명서에서 [Storage types & views]를 참조하세요.

Hudi는 MoR용 Hive 메타스토어에 테이블을 두 개 만듭니다. 즉, 사용자가 지정한 이름을 가진 테이블(읽기 최적화 보기)과 `_rt`가 추가된 동일한 이름을 가진 테이블(실시간 보기)입니다. 두 테이블을 모두 쿼리할 수 있습니다.

Hudi 데이터 세트를 메타스토어에 등록

Hudi 테이블을 Hive 메타스토dj에 등록할 때 기타 모든 테이블처럼 Hive, Spark SQL 또는 Presto를 사용하여 Hudi 테이블을 쿼리할 수 있습니다. 또한 Glue 데이터 카탈로그를 AWS 메타스토어로 사용하도록 Hive 및 Spark를 구성하여 Hudi를 AWS Glue와 통합할 수 있습니다. MoR 테이블의 경우 Hudi는 메타스토어에 두 개의 테이블로 데이터 세트를 등록합니다. 즉, 사용자가 지정한 이름을 가진 테이블(읽기 최적화 보기)과 `_rt`가 추가된 동일한 이름을 가진 테이블(실시간 보기)입니다.

`HIVE_SYNC_ENABLED_OPT_KEY` 옵션을 `"true"`로 설정하고 다른 필수 속성을 제공하여 Hudi 데이터 세트를 만드는 데 Spark를 사용하는 경우 Hudi 테이블을 Hive 메타스토어에 등록합니다. 자세한 내용은 Hudi 데이터 세트 작업 단원을 참조하십시오. 또한 hive_sync_tool 명령줄 유틸리티를 사용하여 Hudi 데이터 세트를 메타스토어의 테이블로 별도로 등록할 수 있습니다.

Amazon EMR 기반 Hudi 사용 시 고려 사항 및 제한 사항

- 레코드 키 필드의 경우 null이거나 비워둘 수 없음 - 레코드 키 필드로 지정하는 필드는 `null`이거나 빈 값일 수 없습니다.
- 업서트 및 삽입 시 기본적으로 업데이트되는 스키마 - Hudi는 입력 DataFrame과 기존 Hudi 데이터 세트를 병합하여 업데이트된 새 데이터 세트를 생성하는 방법을 결정

하는 인터페이스 HoodieRecordPayload를 제공합니다. Hudi는 입력 DataFrame에 지정된 대로 기존 레코드를 덮어쓰고 스키마를 업데이트하는 이 클래스의 기본 구현(OverwriteWithLatestAvroPayload)을 제공합니다. 병합 및 부분 업데이트 구현을 위해 이 논리를 사용자 지정하려면 DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY 파라미터를 사용하여 HoodieRecordPayload 인터페이스 구현을 제공할 수 있습니다.

- 삭제 시 스키마 필요 - 삭제할 때는 레코드 키, 파티션 키 및 사전 결합 키 필드를 지정해야 합니다. 다른 열은 null이거나 비워둘 수 있지만 전체 스키마가 필요합니다.
- MoR 테이블 제한 - MoR 테이블은 세이브 포인트 기능을 지원하지 않습니다. Spark SQL, Presto 또는 Hive의 읽기 최적화 보기 또는 실시간 보기(*tableName*_rt)를 사용하여 MoR 테이블을 쿼리할 수 있습니다. 읽기 최적화 보기를 사용하면 기본 파일 데이터만 노출되고 기본 및 로그 데이터의 병합된 보기는 노출되지 않습니다.
- Hive
  - 테이블을 Hive 메타스토어에 등록하는 경우 Hudi는 Hive Thrift 서버가 기본 포트인 10000에서 실행될 것으로 예상합니다. 이 포트를 사용자 지정 포트로 재정의하는 경우 다음 예와 같이 HIVE_URL_OPT_KEY 옵션을 전달합니다.

```
.option(DataSourceWriteOptions.HIVE_URL_OPT_KEY, "jdbc:hive2://localhost:override-port-number
```

  - Spark의 timestamp 데이터 유형은 Hive의 long 유형이 아닌 Hive의 timestamp 데이터 유형으로 등록됩니다.
- Presto
  - Presto는 0.6.0 미만의 Hudi 버전에서 MoR 실시간 테이블 읽기를 지원하지 않습니다.
  - Presto는 스냅샷 쿼리만 지원합니다.
  - Presto에서 Hudi 데이터 세트 열을 올바르게 해석하려면 hive.parquet_use_column_names 값을 true로 설정합니다.
    - 세션의 값을 설정하려면 Presto 셸에서 다음 명령을 실행합니다.

```
set session hive.parquet_use_column_names=true
```

    - 클러스터 수준에서 값을 설정하려면 다음 예와 같이 presto-connector-hive 구성 분류를 사용하여 hive.parquet.use_column_names을(를) true로 설정합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

```
[
  {
```

```
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.parquet.use-column-names": "true"
    }
  }
]
```

- HBase 인덱스
  - Hudi를 빌드하는 데 사용된 HBase 버전은 EMR 릴리스 안내서에 나온 버전과 다를 수 있습니다. Spark 세션에 올바른 종속 항목을 가져오려면 다음 명령을 실행합니다.

```
spark-shell \
--jars /usr/lib/spark/external/lib/spark-avro.jar,/usr/lib/hudi/cli/lib/*.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false"
```

- 최상의 성능을 위한 설정 - EMR 7.3+/Hudi 0.15+의 경우 고객은 Kryo 직렬화 오버헤드를 줄이기 위해이 구성을 설정하는 것이 좋습니다.

```
--conf 'spark.kryo.registrator=org.apache.spark.HoodieKryoRegistrar'
```

> **Note**
>
> EMR Serverless에서 세분화된 액세스 제어(FGAC)를 사용하는 경우 사용자가 KryoSerializer 대신 JavaSerializer를 사용해야 하므로이 구성은 필요하지 않습니다. KryoSerializer

Hudi가 설치된 클러스터 생성

Amazon EMR 릴리스 버전 5.28.0 이상에서 EMR은 Spark, Hive 또는 Presto가 설치된 경우 기본적으로 Hudi 구성 요소를 설치합니다. Amazon EMR 기반 Hudi를 사용하려면 다음 애플리케이션 중 하나 이상 설치된 클러스터를 생성합니다.

- Hadoop
- Hive
- Spark
- Presto

- Flink

AWS CLI, 또는 Amazon EMR API AWS Management Console를 사용하여 클러스터를 생성할 수 있습니다.

를 사용하여 Hudi로 클러스터를 생성하려면 AWS Management Console

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. 소프트웨어 구성에서 릴리스로 emr-5.28.0 이상을 선택하고 클러스터에 필요한 다른 애플리케이션과 함께 Hadoop, Hive, Spark, Presto, Tez를 선택합니다.
4. 애플리케이션에 필요한 기타 옵션을 구성한 후 [Next]를 선택합니다.
5. 하드웨어 및 일반 클러스터 설정에 대한 옵션을 원하는 대로 구성합니다.
6. 보안 옵션의 경우 SSH를 사용하여 마스터 노드 명령줄에 연결하는 데 사용할 수 있는 EC2 키 페어를 선택하는 것이 좋습니다. 이렇게 하면 이 안내서에 설명된 Spark 쉘 명령, Hive CLI 명령 및 Hudi CLI 명령을 실행할 수 있습니다.
7. 기타 보안 옵션을 원하는 대로 선택하고 클러스터 생성을 선택합니다.

Hudi 데이터 세트 작업

Hudi는 Spark를 통해 Hudi 데이터 세트에서의 데이터 삽입, 업데이트 및 삭제를 지원합니다. 자세한 내용은 Apache Hudi 설명서에서 Writing Hudi tables를 참조하세요.

다음 예제에서는 대화식 Spark 쉘을 시작하거나 Spark 제출을 사용하거나 Amazon EMR에서 Hudi를 작업하기 위해 Amazon EMR Notebooks를 사용하는 방법을 보여줍니다. 또한 Hudi DeltaStreamer 유틸리티 또는 기타 도구를 사용하여 데이터 세트에 쓸 수도 있습니다. 이 섹션에 소개된 예에서는 SSH를 기본 hadoop 사용자로 사용하여 마스터 노드에 연결되어 있는 동안 Spark 셸을 사용하여 데이터 세트로 작업하는 방법을 보여줍니다.

Amazon EMR 6.7 이상을 사용하여 Spark 쉘 시작

Amazon EMR 6.7.0 이상을 사용하여 `spark-shell`, `spark-submit` 또는 `spark-sql`을 실행하는 경우 다음 명령을 전달합니다.

> **Note**
>
> Amazon EMR 6.7.0은 Apache Hudi 0.11.0-amzn-0을 사용하며, 이는 이전 Hudi 버전에 비해 크게 개선되었습니다. 자세한 내용은 Apache Hudi 0.11.0 Migration Guide를 참조하세요. 이 탭의 예제에는 이러한 변경 사항이 반영되어 있습니다.

프라이머리 노드에서 Spark 셸을 여는 방법

1. SSH를 사용하여 프라이머리 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. Spark 셸을 시작하려면 다음 명령을 입력합니다. PySpark 셸을 사용하려면 *spark-shell*을 *pyspark*로 바꿉니다.

```
spark-shell --jars /usr/lib/hudi/hudi-spark-bundle.jar \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.hudi.catalog.HoodieCatalog"
 \
--conf "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
```

Amazon EMR 6.6 이하를 사용하여 Spark 셸 시작

Amazon EMR 6.6.x 이하를 사용하여 `spark-shell`, `spark-submit` 또는 `spark-sql`을 실행하는 경우 다음 명령을 전달합니다.

> **Note**
>
> - Amazon EMR 6.2 및 5.31 이상(Hudi 0.6.x 이상)에서는 구성에서 `spark-avro.jar`을 생략할 수 있습니다.
> - Amazon EMR 6.5 및 5.35 이상(Hudi 0.9.x 이상)에서는 구성에서 `spark.sql.hive.convertMetastoreParquet=false`를 생략할 수 있습니다.
> - Amazon EMR 6.6 및 5.36 이상(Hudi 0.10.x 이상)에서는 Version: 0.10.0 Spark Guide에 설명된 대로 `HoodieSparkSessionExtension` 구성을 포함해야 합니다.

```
--conf
 "spark.sql.extensions=org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
 \
```

프라이머리 노드에서 Spark 셸을 여는 방법

1. SSH를 사용하여 프라이머리 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. Spark 셸을 시작하려면 다음 명령을 입력합니다. PySpark 셸을 사용하려면 *spark-shell*을 *pyspark*로 바꿉니다.

```
spark-shell \
--conf "spark.serializer=org.apache.spark.serializer.KryoSerializer" \
--conf "spark.sql.hive.convertMetastoreParquet=false" \
--jars /usr/lib/hudi/hudi-spark-bundle.jar,/usr/lib/spark/external/lib/spark-
avro.jar
```

Amazon EMR 6.7 이상을 사용하는 Amazon EMR Notebooks에서 Hudi 사용

Amazon EMR Notebooks에서 Hudi를 사용하려면 먼저 로컬 파일 시스템의 Hudi jar 파일을 노트북 클러스터의 프라이머리 노드에 있는 HDFS로 복사해야 합니다. 그런 다음, 노트북 편집기를 이용해 Hudi를 사용하도록 EMR 노트북을 구성합니다.

Amazon EMR Notebooks에서 Hudi를 사용하는 방법

1. Amazon EMR Notebooks에 대한 클러스터를 생성하고 시작합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 노트북에 대한 Amazon EMR 클러스터 생성을 참조하세요.
2. SSH를 사용하여 클러스터의 마스터 노드에 연결하고 다음 예와 같이 로컬 파일 시스템의 jar 파일을 HDFS로 복사합니다. 이 예에서는 파일 관리의 명확성을 위해 HDFS에 디렉터리를 만듭니다. 원할 경우 HDFS에서 자체 대상을 선택할 수 있습니다.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-
spark-bundle.jar
```

3. 노트북 편집기를 열고 다음 예의 코드를 입력하고 실행합니다.

```
%%configure
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar",
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.catalog.spark_catalog":
 "org.apache.spark.sql.hudi.catalog.HoodieCatalog",

 "spark.sql.extensions":"org.apache.spark.sql.hudi.HoodieSparkSessionExtension"
          }}
```

Amazon EMR 6.6 이하를 사용하는 Amazon EMR Notebooks에서 Hudi 사용

Amazon EMR Notebooks에서 Hudi를 사용하려면 먼저 로컬 파일 시스템의 Hudi jar 파일을 노트북 클러스터의 프라이머리 노드에 있는 HDFS로 복사해야 합니다. 그런 다음, 노트북 편집기를 이용해 Hudi를 사용하도록 EMR 노트북을 구성합니다.

Amazon EMR Notebooks에서 Hudi를 사용하는 방법

1. Amazon EMR Notebooks에 대한 클러스터를 생성하고 시작합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 노트북에 대한 Amazon EMR 클러스터 생성을 참조하세요.
2. SSH를 사용하여 클러스터의 마스터 노드에 연결하고 다음 예와 같이 로컬 파일 시스템의 jar 파일을 HDFS로 복사합니다. 이 예에서는 파일 관리의 명확성을 위해 HDFS에 디렉터리를 만듭니다. 원할 경우 HDFS에서 자체 대상을 선택할 수 있습니다.

```
hdfs dfs -mkdir -p /apps/hudi/lib
```

```
hdfs dfs -copyFromLocal /usr/lib/hudi/hudi-spark-bundle.jar /apps/hudi/lib/hudi-
spark-bundle.jar
```

```
hdfs dfs -copyFromLocal /usr/lib/spark/external/lib/spark-avro.jar /apps/hudi/lib/
spark-avro.jar
```

3. 노트북 편집기를 열고 다음 예의 코드를 입력하고 실행합니다.

```
{ "conf": {
            "spark.jars":"hdfs:///apps/hudi/lib/hudi-spark-bundle.jar,hdfs:///apps/
hudi/lib/spark-avro.jar",
```

```
            "spark.serializer":"org.apache.spark.serializer.KryoSerializer",
            "spark.sql.hive.convertMetastoreParquet":"false"
          }}
```

Hudi용 Spark 세션 초기화

Scala를 사용할 때는 Spark 세션에서 다음 클래스를 가져와야 합니다. Spark 세션당 한 번 이 작업을 수행하면 됩니다.

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._
import org.apache.hudi.DataSourceWriteOptions
import org.apache.hudi.DataSourceReadOptions
import org.apache.hudi.config.HoodieWriteConfig
import org.apache.hudi.hive.MultiPartKeysValueExtractor
import org.apache.hudi.hive.HiveSyncConfig
import org.apache.hudi.sync.common.HoodieSyncConfig
```

Hudi 데이터 세트에 쓰기

다음 예제에서는 DataFrame을 생성하고 Hudi 데이터 세트로 쓰는 방법을 보여줍니다.

> **Note**
>
> 코드 샘플을 Spark 셸에 붙여넣으려면 프롬프트에 **`:paste`**를 입력하고 예제를 붙여넣은 다음 **`CTRL`** + **`D`**를 누릅니다.

Hudi 데이터 세트에 DataFrame을 쓸 때마다 `DataSourceWriteOptions`을 지정해야 합니다. 이러한 옵션 대부분이 쓰기 작업 간에 동일할 수 있습니다. 다음 예에서는 *`hudiOptions`* 변수를 사용하여 일반 옵션을 지정합니다. 이 변수는 그 다음 예제에서도 사용합니다.

Amazon EMR 6.7 이상에서 Scala를 사용하여 작성

> **Note**
>
> Amazon EMR 6.7.0은 [Apache Hudi](https://hudi.apache.org/) 0.11.0-amzn-0을 사용하며, 이는 이전 Hudi 버전에 비해 크게 개선되었습니다. 자세한 내용은 [Apache Hudi 0.11.0 Migration Guide](https://hudi.apache.org/)를 참조하세요. 이 탭의 예제에는 이러한 변경 사항이 반영되어 있습니다.

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TBL_NAME.key -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE.key -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  HoodieSyncConfig.META_SYNC_PARTITION_EXTRACTOR_CLASS.key ->
 "org.apache.hudi.hive.MultiPartKeysValueExtractor",
  HoodieSyncConfig.META_SYNC_ENABLED.key -> "true",
  HiveSyncConfig.HIVE_SYNC_MODE.key -> "hms",
  HoodieSyncConfig.META_SYNC_TABLE_NAME.key -> "tableName",
  HoodieSyncConfig.META_SYNC_PARTITION_FIELDS.key -> "creation_date"
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,"insert")
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Amazon EMR 6.6 이하에서 Scala를 사용하여 작성

```
// Create a DataFrame
val inputDF = Seq(
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
```

```
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
 ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
 ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z")
 ).toDF("id", "creation_date", "last_update_time")

//Specify common DataSourceWriteOptions in the single hudiOptions variable
val hudiOptions = Map[String,String](
  HoodieWriteConfig.TABLE_NAME -> "tableName",
  DataSourceWriteOptions.TABLE_TYPE_OPT_KEY -> "COPY_ON_WRITE",
  DataSourceWriteOptions.RECORDKEY_FIELD_OPT_KEY -> "id",
  DataSourceWriteOptions.PARTITIONPATH_FIELD_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.PRECOMBINE_FIELD_OPT_KEY -> "last_update_time",
  DataSourceWriteOptions.HIVE_SYNC_ENABLED_OPT_KEY -> "true",
  DataSourceWriteOptions.HIVE_TABLE_OPT_KEY -> "tableName",
  DataSourceWriteOptions.HIVE_PARTITION_FIELDS_OPT_KEY -> "creation_date",
  DataSourceWriteOptions.HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY ->
 classOf[MultiPartKeysValueExtractor].getName
)

// Write the DataFrame as a Hudi dataset
(inputDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.INSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Overwrite)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

PySpark를 사용하여 작성

```
# Create a DataFrame
inputDF = spark.createDataFrame(
    [
        ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
        ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
        ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
        ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z"),
        ("104", "2015-01-02", "2015-01-01T12:15:00.512679Z"),
        ("105", "2015-01-02", "2015-01-01T13:51:42.248818Z"),
    ],
    ["id", "creation_date", "last_update_time"]
)
```

```
# Specify common DataSourceWriteOptions in the single hudiOptions variable
hudiOptions = {
'hoodie.table.name': 'tableName',
'hoodie.datasource.write.recordkey.field': 'id',
'hoodie.datasource.write.partitionpath.field': 'creation_date',
'hoodie.datasource.write.precombine.field': 'last_update_time',
'hoodie.datasource.hive_sync.enable': 'true',
'hoodie.datasource.hive_sync.table': 'tableName',
'hoodie.datasource.hive_sync.partition_fields': 'creation_date',
'hoodie.datasource.hive_sync.partition_extractor_class':
 'org.apache.hudi.hive.MultiPartKeysValueExtractor'
}

# Write a DataFrame as a Hudi dataset
inputDF.write \
.format('org.apache.hudi') \
.option('hoodie.datasource.write.operation', 'insert') \
.options(**hudiOptions) \
.mode('overwrite') \
.save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

> **Note**
>
> 코드 예제 및 알림에서 Hudi 대신 'hoodie'가 표시될 수 있습니다. Hudi 코드베이스는 예전의 'hoodie' 철자를 널리 사용합니다.

Hudi에 대한 DataSourceWriteOptions 참조

옵션	설명
TABLE_NAME	데이터 세트를 등록할 테이블 이름입니다.
TABLE_TYPE_OPT_KEY	선택 사항. 데이터 세트를 `"COPY_ON_WRITE"` 또는 `"MERGE_ON_READ"` 로 생성할지 여부를 지정합니다. 기본값은 `"COPY_ON_WRITE"` 입니다.
RECORDKEY_FIELD_OPT_KEY	값의 레코드 키 필드는 `recordKey` 의 `HoodieKey` 구성 요소로 사용됩니다. 실제 값은 필드 값의 `.toString()` 을 호출하여 얻을

옵션	설명
	수 있습니다. 중첩 필드는 점 표기법을 사용하여 지정할 수 있습니다(예: `a.b.c`).
PARTITIONPATH_FIELD_OPT_KEY	값의 파티션 경로 필드는 `HoodieKey` 의 `partitionPath` 구성 요소로 사용됩니다. 실제 값은 필드 값의 `.toString()` 을 호출하여 얻을 수 있습니다.
PRECOMBINE_FIELD_OPT_KEY	실제로 쓰기 전에 미리 결합하는 데 사용되는 필드입니다. 두 레코드에 동일한 키 값이 있는 경우 Hudi는 `Object.compareTo(..)` 에 의해 결정된 사전 결합 필드에 대해 가장 큰 값을 가진 레코드를 선택합니다.

다음 옵션은 메타스토어에 Hudi 데이터 세트 테이블을 등록하는 데에만 필요합니다. Hudi 데이터 세트를 Hive 메타스토어의 테이블로 등록하지 않은 경우 이러한 옵션이 필요하지 않습니다.

Hive에 대한 DataSourceWriteOptions 참조

옵션	설명
HIVE_DATABASE_OPT_KEY	동기화할 Hive 데이터베이스입니다. 기본값은 `"default"` 입니다.
HIVE_PARTITION_EXTRACTOR_CLASS_OPT_KEY	파티션 필드 값을 Hive 파티션 열로 추출하는 데 사용되는 클래스입니다.
HIVE_PARTITION_FIELDS_OPT_KEY	Hive 파티션 열을 결정하는 데 사용할 데이터 세트의 필드입니다.
HIVE_SYNC_ENABLED_OPT_KEY	`"true"`로 설정하면 데이터 세트를 Apache Hive 메타스토어에 등록합니다. 기본값은 `"false"`입니다.
HIVE_TABLE_OPT_KEY	필수 사항입니다. 동기화할 Hive의 테이블의 이름입니다. 예를 들어 `"my_hudi_table_cow"` 입니다.

옵션	설명
HIVE_USER_OPT_KEY	선택 사항. 동기화할 때 사용할 Hive 사용자 이름입니다. 예를 들어 "hadoop"입니다.
HIVE_PASS_OPT_KEY	선택 사항. HIVE_USER_OPT_KEY 에 의해 지정된 사용자의 Hive 암호입니다.
HIVE_URL_OPT_KEY	Hive 메타스토어 URL입니다.

데이터 업서트

다음 예제에서는 DataFrame를 작성하여 데이터를 업서트하는 방법을 보여줍니다. 이전의 삽입 예제와 달리 OPERATION_OPT_KEY 값은 UPSERT_OPERATION_OPT_VAL로 설정됩니다. 또한 레코드가 추가되어야 함을 나타내기 위해 .mode(SaveMode.Append)가 지정됩니다.

Amazon EMR 6.7 이상에서 Scala를 사용하여 업서트

Note

Amazon EMR 6.7.0은 Apache Hudi 0.11.0-amzn-0을 사용하며, 이는 이전 Hudi 버전에 비해 크게 개선되었습니다. 자세한 내용은 Apache Hudi 0.11.0 Migration Guide를 참조하세요. 이 탭의 예제에는 이러한 변경 사항이 반영되어 있습니다.

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "upsert")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Amazon EMR 6.6 이하에서 Scala를 사용하여 업서트

```
// Create a new DataFrame from the first row of inputDF with a different creation_date
 value
val updateDF = inputDF.limit(1).withColumn("creation_date", lit("new_value"))

(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .options(hudiOptions)
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

PySpark를 사용하여 업서트

```
from pyspark.sql.functions import lit

# Create a new DataFrame from the first row of inputDF with a different creation_date
 value
updateDF = inputDF.limit(1).withColumn('creation_date', lit('new_value'))

updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

레코드 삭제

레코드를 하드 삭제하려면 빈 페이로드를 업서트할 수 있습니다. 이 경우 PAYLOAD_CLASS_OPT_KEY 옵션은 EmptyHoodieRecordPayload 클래스를 지정합니다. 이 예제에서는 동일한 레코드를 지정하도록 업서트 예제에서 사용된 동일한 DataFrame(updateDF)을 사용합니다.

Amazon EMR 6.7 이상에서 Scala를 사용하여 삭제

Note

Amazon EMR 6.7.0은 Apache Hudi 0.11.0-amzn-0을 사용하며, 이는 이전 Hudi 버전에 비해 크게 개선되었습니다. 자세한 내용은 Apache Hudi 0.11.0 Migration Guide를 참조하세요. 이 탭의 예제에는 이러한 변경 사항이 반영되어 있습니다.

```
(updateDF.write
    .format("hudi")
    .options(hudiOptions)
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY, "delete")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

Amazon EMR 6.6 이하에서 Scala를 사용하여 삭제

```
(updateDF.write
    .format("org.apache.hudi")
    .option(DataSourceWriteOptions.OPERATION_OPT_KEY,
 DataSourceWriteOptions.UPSERT_OPERATION_OPT_VAL)
    .option(DataSourceWriteOptions.PAYLOAD_CLASS_OPT_KEY,
 "org.apache.hudi.common.model.EmptyHoodieRecordPayload")
    .mode(SaveMode.Append)
    .save("s3://amzn-s3-demo-bucket/myhudidataset/"))
```

PySpark를 사용하여 삭제

```
updateDF.write \
    .format('org.apache.hudi') \
    .option('hoodie.datasource.write.operation', 'upsert') \
    .option('hoodie.datasource.write.payload.class',
 'org.apache.hudi.common.model.EmptyHoodieRecordPayload') \
    .options(**hudiOptions) \
    .mode('append') \
    .save('s3://amzn-s3-demo-bucket/myhudidataset/')
```

제출한 데이터 세트의 모든 레코드를 제거하도록 `OPERATION_OPT_KEY` 를 `DELETE_OPERATION_OPT_VAL`로 설정하여 데이터를 하드 삭제할 수도 있습니다. 소프트 삭제 수행

지침 및 Hudi 테이블에 저장된 데이터 삭제에 대한 자세한 내용은 Apache Hudi 설명서에서 Deletes를 참조하세요.

Hudi 데이터 세트에서 읽기

Hudi는 현재 특정 시점에서 데이터를 검색하기 위해 기본적으로 스냅샷 쿼리를 수행합니다. 다음은 Hudi 데이터 세트에 쓰기에서 S3에 기록된 데이터 세트를 쿼리하는 예제입니다. *s3://amzn-s3-demo-bucket/myhudidataset*를 테이블 경로로 바꾸고 각 파티션 수준에 대한 와일드카드 별표와 하나의 추가 별표를 추가합니다. 이 예제에서는 파티션 수준이 하나이므로 와일드카드 기호를 두 개 추가했습니다.

Amazon EMR 6.7 이상에서 Scala를 사용하여 읽기

> **Note**
>
> Amazon EMR 6.7.0은 Apache Hudi 0.11.0-amzn-0을 사용하며, 이는 이전 Hudi 버전에 비해 크게 개선되었습니다. 자세한 내용은 Apache Hudi 0.11.0 Migration Guide를 참조하세요. 이 탭의 예제에는 이러한 변경 사항이 반영되어 있습니다.

```
val snapshotQueryDF = spark.read
    .format("hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset")
    .show()
```

Amazon EMR 6.6 이하에서 Scala를 사용하여 읽기

```
(val snapshotQueryDF = spark.read
    .format("org.apache.hudi")
    .load("s3://amzn-s3-demo-bucket/myhudidataset" + "/*/*"))

snapshotQueryDF.show()
```

PySpark를 사용하여 읽기

```
snapshotQueryDF = spark.read \
    .format('org.apache.hudi') \
    .load('s3://amzn-s3-demo-bucket/myhudidataset' + '/*/*')
```

```
snapshotQueryDF.show()
```

증분 쿼리

Hudi에서 증분 쿼리를 수행하여 지정된 커밋 타임스탬프 이후 변경된 레코드 스트림을 가져올 수도 있습니다. 이렇게 하려면 QUERY_TYPE_OPT_KEY 필드를 QUERY_TYPE_INCREMENTAL_OPT_VAL로 설정합니다. 그런 다음 BEGIN_INSTANTTIME_OPT_KEY에 대한 값을 추가하여 지정된 시간 이후에 기록된 모든 레코드를 가져옵니다. 증분 쿼리는 변경된 레코드만 처리하므로 배치 쿼리보다 보통 10배 더 효율적입니다.

증분 쿼리를 수행할 때는 스냅샷 쿼리에 와일드카드 별표를 사용하지 않고 루트(기본) 테이블 경로를 사용합니다.

> **Note**
>
> Presto는 증분 쿼리를 지원하지 않습니다.

Scala를 사용한 증분 쿼리

```
(val incQueryDF = spark.read
    .format("org.apache.hudi")
    .option(DataSourceReadOptions.QUERY_TYPE_OPT_KEY,
 DataSourceReadOptions.QUERY_TYPE_INCREMENTAL_OPT_VAL)
    .option(DataSourceReadOptions.BEGIN_INSTANTTIME_OPT_KEY, <beginInstantTime>)
    .load("s3://amzn-s3-demo-bucket/myhudidataset" ))

incQueryDF.show()
```

PySpark를 사용한 증분 쿼리

```
readOptions = {
  'hoodie.datasource.query.type': 'incremental',
  'hoodie.datasource.read.begin.instanttime': <beginInstantTime>,
}

incQueryDF = spark.read \
    .format('org.apache.hudi') \
    .options(**readOptions) \
    .load('s3://amzn-s3-demo-bucket/myhudidataset')
```

```
incQueryDF.show()
```

Hudi 데이터 세트에서 읽는 방법에 대한 자세한 내용은 Apache Hudi 설명서에서 Querying Hudi tables를 참조하세요.

Hudi CLI 사용

Hudi CLI를 사용하여 Hudi 데이터 세트를 관리하고 커밋, 파일 시스템, 통계 등에 대한 정보를 볼 수 있습니다. 또한 CLI를 사용하여 수동으로 압축을 수행하거나, 압축을 예약하거나, 예약된 압축을 취소할 수도 있습니다. 자세한 내용은 Apache Hudi 설명서에서 Interacting via CLI를 참조하세요.

Hudi CLI를 시작하고 데이터 세트에 연결하는 방법

1. SSH를 사용하여 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. 명령줄에 `/usr/lib/hudi/cli/bin/hudi-cli.sh`를 입력합니다.

   명령 프롬프트가 이렇게 바뀝니다. `hudi->`
3. 다음 명령을 사용하여 데이터세트에 연결합니다. *s3://amzn-s3-demo-bucket/myhudidataset*를 작업하려는 데이터세트의 경로로 바꿉니다. 여기서 사용하는 값은 이전 예에서 설정한 값과 동일합니다.

   ```
   connect --path s3://amzn-s3-demo-bucket/myhudidataset
   ```

   다음 예와 같이 연결한 데이터 세트를 포함하도록 명령 프롬프트가 변경됩니다.

   ```
   hudi:myhudidataset->
   ```

기본적으로 Amazon EMR 릴리스 `hudi-cli.sh` 7.3.0~7.8.0의 스크립트는를 사용합니다`hudi-cli-bundle.jar`. 문제가 발생하면 다음 명령을 사용하여 클래식 Hudi CLI로 다시 전환할 수 있습니다.

```
/usr/lib/hudi/cli/bin/hudi-cli.sh --cliBundle false
```

이 명령은 `hudi-cli.sh` 스크립트를 실행하고, `--cliBundle` 플래그를 설정하며, 번들에 포함된 JAR 대신 개별 JAR 파일을 사용하도록 CLI에 지시합니다. 기본적으로 `--cliBundle`은 true로 설정되어 있습니다. 즉, CLI는 대신 번들에 포함된 JAR을 사용합니다.

Amazon EMR 7.9.0 이상 릴리스 사용

> **Note**
>
> EMR 릴리스 7.9.0 이상에서는 hudi-cli.sh 스크립트가 더 이상 사용되지 않습니다. Amazon EMR 릴리스 7.9.0 이상에서는 hudi-cli-bundle.jar을 사용합니다.

Hudi CLI를 시작하고 데이터 세트에 연결하려면:

1. SSH를 사용하여 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [SSH를 사용하여 프라이머리 노드에 연결](#)을 참조하세요.
2. 명령줄에서 /usr/lib/hudi/cli-bundle/bin/hudi-cli-with-bundle.sh를 입력하거나 hudi-cli-with-bundle 또는 >hudi-cli를 입력하면 됩니다.

   명령 프롬프트가 hudi- >로 변경됩니다.
3. 다음 명령을 사용하여 데이터세트에 연결합니다. s3://amzn-s3-demo-bucket/myhudidataset를 작업하려는 데이터세트의 경로로 바꿉니다. 여기서 사용하는 값은 이전 예에서 설정한 값과 동일합니다.

```
connect --path s3://amzn-s3-demo-bucket/myhudidataset
```

4. 다음 예와 같이 연결한 데이터 세트를 포함하도록 명령 프롬프트가 변경됩니다.

```
hudi:myhudidataset->
```

Hudi 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Hudi의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Hudi 버전 정보

Amazon EMR 릴리스 레이블	Hudi 버전	Hudi와 함께 설치된 구성 요소
emr-7.9.0	0.15.0-amzn-6	Not available.

Amazon EMR 릴리스 레이블	Hudi 버전	Hudi와 함께 설치된 구성 요소
emr-7.8.0	0.15.0-amzn-5	Not available.
emr-7.7.0	0.15.0-amzn-4	Not available.
emr-7.6.0	0.15.0-amzn-3	Not available.
emr-7.5.0	0.15.0-amzn-2	Not available.
emr-7.4.0	0.15.0-amzn-1	Not available.
emr-7.3.0	0.15.0-amzn-0	Not available.
emr-7.2.0	0.14.1-amzn-1	Not available.
emr-5.36.2	0.10.1-amzn-1	Not available.
emr-7.1.0	0.14.1-amzn-0	Not available.
emr-7.0.0	0.14.0-amzn-1	Not available.
emr-6.15.0	0.14.0-amzn-0	Not available.
emr-6.14.0	0.13.1-amzn-2	Not available.
emr-6.13.0	0.13.1-amzn-1	Not available.
emr-6.12.0	0.13.1-amzn-0	Not available.
emr-6.11.1	0.13.0-amzn-0	Not available.
emr-6.11.0	0.13.0-amzn-0	Not available.
emr-6.10.1	0.12.2-amzn-0	Not available.
emr-6.10.0	0.12.2-amzn-0	Not available.
emr-6.9.1	0.12.1-amzn-0	Not available.
emr-6.9.0	0.12.1-amzn-0	Not available.

Amazon EMR 릴리스 레이블	Hudi 버전	Hudi와 함께 설치된 구성 요소
emr-6.8.1	0.11.1-amzn-0	Not available.
emr-6.8.0	0.11.1-amzn-0	Not available.
emr-6.7.0	0.11.0-amzn-0	Not available.
emr-5.36.1	0.10.1-amzn-1	Not available.
emr-5.36.0	0.10.1-amzn-1	Not available.
emr-6.6.0	0.10.1-amzn-0	Not available.
emr-5.35.0	0.9.0-amzn-2	Not available.
emr-6.5.0	0.9.0-amzn-1	Not available.
emr-6.4.0	0.8.0-amzn-0	Not available.
emr-6.3.1	0.7.0-amzn-0	Not available.
emr-6.3.0	0.7.0-amzn-0	Not available.
emr-6.2.1	0.6.0-amzn-1	Not available.
emr-6.2.0	0.6.0-amzn-1	Not available.
emr-6.1.1	0.5.2-incubating-amzn-2	Not available.
emr-6.1.0	0.5.2-incubating-amzn-2	Not available.
emr-6.0.1	0.5.0-incubating-amzn-1	Not available.
emr-6.0.0	0.5.0-incubating-amzn-1	Not available.
emr-5.34.0	0.9.0-amzn-0	Not available.
emr-5.33.1	0.7.0-amzn-1	Not available.
emr-5.33.0	0.7.0-amzn-1	Not available.

Amazon EMR 릴리스 레이블	Hudi 버전	Hudi와 함께 설치된 구성 요소
emr-5.32.1	0.6.0-amzn-0	Not available.
emr-5.32.0	0.6.0-amzn-0	Not available.
emr-5.31.1	0.6.0-amzn-0	Not available.
emr-5.31.0	0.6.0-amzn-0	Not available.
emr-5.30.2	0.5.2-incubating	Not available.
emr-5.30.1	0.5.2-incubating	Not available.
emr-5.30.0	0.5.2-incubating	Not available.
emr-5.29.0	0.5.0-incubating	Not available.
emr-5.28.1	0.5.0-incubating	Not available.
emr-5.28.0	0.5.0-incubating	Not available.

Hue

Hue(Hadoop 사용자 환경)는 Amazon EMR 및 Apache Hadoop과 함께 사용하는 오픈 소스 웹 기반 그래픽 사용자 인터페이스입니다 . Hue는 여러 가지 다른 Hadoop 에코시스템 프로젝트를 구성 가능한 인터페이스로 함께 그룹화합니다. Amazon EMR은 Amazon EMR의 Hue에 관련된 사용자 지정도 추가했습니다. Hue는 클러스터에서 실행되는 애플리케이션의 프런트 엔드 역할을 하여 보다 익숙한 인터페이스를 통해 애플리케이션과 상호 작용할 수 있게 해줍니다. Hue에서 Hive 및 Pig 편집기와 같은 애플리케이션을 사용하면 클러스터에 로그인하여 각 애플리케이션의 셸을 사용하여 대화식으로 스크립트를 실행할 필요가 없습니다. 클러스터 시작 후 Hue 또는 유사한 인터페이스를 사용하여 애플리케이션과 완전히 상호 작용할 수 있습니다. Hue에 대한 자세한 내용은 http://gethue.com을 참조하십시오.

Hue는 Amazon EMR 콘솔을 사용하여 클러스터를 시작할 때 기본적으로 설치됩니다. 클러스터를 시작할 때 Amazon EMR 콘솔의 고급 옵션을 사용하거나 AWS CLI에서 `create-cluster`를 사용할 때는 `--applications` 옵션을 명시적으로 지정하고 Hue를 생략하여 Hue를 설치하지 않도록 선택할 수 있습니다.

주제

Hue 버전 정보

7.9.0용 Hue 버전

다음 테이블에는 Amazon EMR이 Hue를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Hue의 버전이 나열되어 있습니다.

이 릴리스에서 Hue와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Hue 버전 정보

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-7.9.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

6.15.0용 Hue 버전

다음 테이블에는 Amazon EMR이 Hue를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Hue의 버전이 나열되어 있습니다.

이 릴리스에서 Hue와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Hue 버전 정보

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.15.0	Hue 4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

5.36.2용 Hue 버전

다음 테이블에는 Amazon EMR이 Hue를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Hue의 버전이 나열되어 있습니다.

이 릴리스에서 Hue와 함께 설치된 구성 요소의 버전은 [릴리스 5.36.2 구성 요소 버전]을 참조하세요.

emr-5.36.2용 Hue 버전 정보

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.36.2	Hue 4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR에서 지원되는 Hue의 기능 및 지원되지 않는 기능

Amazon EMR 7.0.0 이상을 사용하는 경우 Hue에는 Python 3.9 이상이 있어야 합니다. 전송 중 암호화를 사용할 때에는 2048비트 암호 키를 사용해야 합니다.

- Amazon S3 및 Hadoop File System (HDFS) Browser
  - 적절한 권한이 있다면 계정에 속한 휘발성 HDFS 스토리지 및 S3 버킷 간에 데이터를 찾아보고 이동할 수 있습니다.
  - 기본적으로 Hue의 수퍼유저는 Amazon EMR IAM 역할에서 액세스할 수 있는 모든 파일에 액세스할 수 있습니다. 새로 생성된 사용자에게는 Amazon S3 파일 브라우저에 액세스할 수 있는 권한이 자동으로 부여되지 않기 때문에 사용자의 그룹에 대해 활성화된 `filebrowser.s3_access` 권한이 있어야 합니다.
- Hive — 데이터에서 대화형 쿼리를 실행합니다. 이는 프로그래밍 방식 또는 배치 처리된 쿼리를 프로토타입하는 유용한 방법이기도 합니다.

- Pig — 데이터에서 스크립트를 실행하거나 대화형 명령을 실행합니다.
- Oozie — Oozie 워크플로를 생성하고 모니터링합니다.
- Metastore Manager — Hive 메타스토어의 콘텐츠를 보고 조작합니다(가져오기, 생성, 삭제 등).
- 작업 브라우저 — 제출된 Hadoop 작업의 상태를 봅니다.
- 사용자 관리 — 사용자 계정을 관리하고 LDAP 사용자를 Hue와 통합합니다.
- AWS 샘플 - Hue의 애플리케이션을 사용하여 다양한 AWS 서비스의 샘플 데이터를 처리하는 몇 가지 "ready-to-run 완료" 예제가 있습니다. Hue에 로그인하면 샘플이 미리 설치되어 있는 Hue Home 애플리케이션으로 이동됩니다.
- Livy Server는 Amazon EMR 5.9.0 및 이후 버전에서만 지원됩니다.
- Spark용 Hue Notebook을 사용하려면 Livy and Spark와 함께 Hue를 설치해야합니다.
- Trino/Presto — Hue에서는 Trino 및 Presto 데이터 서버 연결을 지원합니다. Amazon EMR 7.0.0 이상에서는 `trino-python-client` 및 `presto-python-client` 커넥터를 사용하여 이러한 연결을 원활하게 만들 수 있습니다. 7.0.0 미만의 릴리스는 JDBC 게이트웨이를 통한 JDBC 커넥터를 사용합니다.
- Hue 대시보드는 지원되지 않습니다.
- PostgreSQL은 지원되지 않습니다.

고려 사항

Amazon EMR에서 Hue를 사용하는 경우 다음과 같은 제한 사항을 고려합니다.

대형 Hue 메타데이터 테이블에서 성능

Hue 메타데이터 데이터베이스가 너무 크면 성능이 저하될 수 있습니다. 테이블 크기를 확인하려면 먼저 SSH를 사용하여 Amazon EMR on Amazon EC2 클러스터의 프라이머리 노드에 연결하고 `sudo mysql -u root` 명령을 실행하여 MySQL CLI를 시작합니다. 테이블 크기를 가져오려면 `SELECT COUNT(*) FROM hue.`*`<table_name>`* 쿼리를 실행합니다. *`<table_name>`*에 대해 가능한 항목은 다음을 참조하세요.

- desktop_document
- desktop_document2
- oozie_job
- beeswax_savedquery

- beeswax_session
- beeswax_queryhistory

이 쿼리를 실행할 때 100,000개가 넘는 수를 반환하는 경우 다음 정리 명령을 실행하여 이전 레코드를 삭제해야 합니다.

```
cd /opt/cloudera/parcels/CDH/lib/hue # Hue home directory ./build/env/bin/hue
 desktop_document_cleanup
```

데이터베이스 정리에 대한 자세한 내용은 Hue 설명서의 reference 페이지를 참조하세요.

Hue는 테이블을 자동으로 정리하지 않지만 Amazon EMR 릴리스 5.12.0 이상에서는 테이블에서 이전 문서를 삭제하는 방법을 제공합니다. 다음 쉘 스크립트를 생성하고 메타데이터 데이터베이스에 문서를 보관하는 최대 일수를 나타내는 정수 파라미터를 사용하여 Amazon EMR 클러스터에서 단계로 실행합니다.

```
#!/bin/bash
if grep isMaster /mnt/var/lib/info/instance.json | grep false;
then
  echo "This is not the primary node; do nothing, exiting"
  exit 0
fi
while [ ! -f /usr/lib/hue/desktop/core/src/desktop/management/commands/
desktop_document_cleanup.py ]
do
  sleep 1
done
sudo systemctl stop hue.service
sudo sed -i 's+  LOG.warn+  # LOG.warn+g' /usr/lib/hue/desktop/core/src/desktop/
management/commands/desktop_document_cleanup.py
sudo /usr/lib/hue/build/env/bin/hue desktop_document_cleanup --keep-days $1
sudo systemctl start hue.service
```

Hue 버전 간 비호환성

여러 Hue 지원 클러스터에서 동일한 Hue 메타데이터 데이터베이스를 사용하는 경우 이러한 클러스터에서 동일한 버전의 Hue를 실행하는 것이 좋습니다. Hue의 버전마다 Hue 메타데이터 데이터베이스에 대해 다른 스키마를 보유할 수 있습니다. 다른 버전에 대해 동일한 데이터베이스를 사용하면 Hue 설치가 실패할 수 있습니다. 예를 들어 4.10.0 및 4.11.0이 설치된 두 클러스터에 대해 동일한 데이터베이스

를 사용하려고 하면 사용자가 4.10.0 클러스터에 로그인하려고 할 때 로그인 오류가 발생할 수 있습니다.

Amazon EMR 릴리스 레이블 목록과 Hue의 해당 설치 버전은 Hue 릴리스 기록을 참조하세요.

Hue 웹 사용자 인터페이스에 연결

Hue 웹 사용자 인터페이스에 연결하는 방법은 클러스터의 프라이머리 노드에 호스팅된 HTTP 인터페이스에 연결하는 방법과 동일합니다. 다음 절차에서는 Hue 사용자 인터페이스에 액세스하는 방법을 설명합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR 클러스터에 호스팅된 웹 인터페이스 보기를 참조하세요.

Hue 웹 사용자 인터페이스를 보는 방법

1. Amazon EMR 관리 안내서에서 동적 포트 전달을 사용하여 프라이머리 노드에 SSH 터널을 설정하려면 다음 지침을 따릅니다.
2. 브라우저에 `http://master public DNS:8888` 주소를 입력하여 Hue 웹 인터페이스를 엽니다. 예를 들어, *master public dns*는 클러스터 프라이머리 노드의 퍼블릭 DNS 이름입니다(예: `ec2-11-22-333-44.compute-1.amazonaws.com`).
3. 처음 로그인한 관리자인 경우 Hue 로그인 화면에서 사용자 이름 및 암호를 입력하여 Hue 수퍼 계정을 생성한 다음 계정 생성을 선택합니다. 그렇지 않으면 사용자 이름과 암호를 입력하고 계정 생성을 선택하거나 관리자가 제공한 자격 증명을 입력합니다.

Amazon RDS에서 원격 데이터베이스와 함께 Hue 사용

기본적으로 Hue 정보 및 쿼리 기록은 마스터 노드의 로컬 MySQL 데이터베이스에 저장됩니다. 또는 Amazon S3에 저장된 구성과 Amazon Relational Database Service(RDS)의 MySQL 데이터베이스를 사용하여 하나 이상의 Hue 지원 클러스터를 생성할 수 있습니다. 이렇게 하면 Amazon EMR 클러스터의 실행 상태를 유지하지 않고도 Hue에서 생성된 사용자 정보 및 쿼리 기록을 유지할 수 있습니다. Amazon S3 서버 측 암호화를 사용하여 구성 파일을 저장하는 것이 좋습니다.

먼저 Hue용 원격 데이터베이스를 생성합니다.

외부 MySQL 데이터베이스를 생성하려면

1. https://console.aws.amazon.com/rds/에서 Amazon RDS 콘솔을 엽니다.
2. DB 인스턴스 시작을 클릭합니다.

3. MySQL을 선택하고 선택을 클릭합니다.
4. 기본으로 선택된 Multi-AZ Deployment and Provisioned IOPS Storage(다중 AZ 배포 및 프로비저닝된 IOPS 스토리지)를 그대로 두고 다음을 클릭합니다.
5. 이스턴스 사양에서 기본값을 그대로 사용하며 설정을 지정하고 다음을 클릭합니다.
6. [Configure Advanced Settings] 페이지에서 올바른 보안 그룹과 데이터베이스 이름을 선택합니다. 사용하는 보안 그룹은 클러스터의 마스터 노드에서 적어도 포트 3306에 대한 수신 TCP 액세스를 허용해야 합니다. 이때 클러스터를 생성하지 않은 경우 모든 호스트를 포트 3306에 연결하도록 허용하고 클러스터를 시작한 후 보안 그룹을 조정할 수 있습니다. DB 인스턴스 시작을 클릭합니다.
7. RDS 대시보드에서 인스턴스를 선택하고 방금 생성한 인스턴스를 선택합니다. 데이터베이스를 사용할 수 있는 경우 dbname, username, password 및 RDS instance hostname을 기록해 둡니다. 클러스터를 생성하고 구성할 때 이 정보를 사용합니다.

를 사용하여 클러스터를 시작할 때 Hue에 대한 외부 MySQL 데이터베이스를 지정하려면 AWS CLI

를 사용하여 클러스터를 시작할 때 Hue용 외부 MySQL 데이터베이스를 지정하려면 구성 객체`hue.ini`로를 구성하기 위해 RDS 인스턴스를 생성할 때 기록해 둔 정보를 AWS CLI사용합니다.

> **Note**
> 동일한 외부 데이터베이스를 사용하는 여러 클러스터를 생성할 수 있지만, 각 클러스터는 쿼리 기록 및 사용자 정보를 공유합니다.

- 를 사용하여 Hue가 설치된 클러스터를 AWS CLI생성하고, 생성한 외부 데이터베이스를 사용하고, 데이터베이스 속성을 지정하는 Hue에 대한 구성 분류가 있는 구성 파일을 참조합니다. 다음은 Hue가 설치된 클러스터를 생성하고 데이터베이스 구성을 지정하는 Amazon S3의 구성 파일(`myConfig.json`)을 참조하는 예제입니다.

  > **Note**
  > 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

  ```
  aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
   Name=Hive \
  --instance-type m5.xlarge --instance-count 3 \
  ```

```
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
myConfig.json --use-default-roles
```

다음은 `myConfig.json` 파일 콘텐츠 예제입니다. *dbname*, *username*, *password* 및 *RDS instance hostname*을 RDS 대시보드에서 이전에 기록한 값으로 대체합니다.

```
[{
  "Classification": "hue-ini",
  "Properties": {},
  "Configurations": [
    {
      "Classification": "desktop",
      "Properties": {},
      "Configurations": [
        {
          "Classification": "database",
          "Properties": {
            "name": "dbname",
            "user": "username",
            "password": "password",
            "host": "RDS instance hostname",
            "port": "3306",
            "engine": "mysql"
          },
          "Configurations": []
        }
      ]
    }
  ]
}]
```

문제 해결

Amazon RDS 장애 조치 시

Hue 데이터베이스 인스턴스가 응답하지 않거나 장애 조치 프로세스 중이어서 쿼리 실행 시 사용자에게 지연이 발생할 수도 있습니다. 이 문제와 관련하여 다음 사항에 유의해야 합니다.

- Amazon RDS 콘솔에 로그인하는 경우 장애 조치 이벤트를 검색할 수 있습니다. 예를 들어 장애 조치가 진행 중이거나 발생한 경우 "Multi-AZ instance failover started" 및 "Multi-AZ instance failover completed." 같은 이벤트가 있는지 살펴봅니다.

- RDS 인스턴스에 대해 장애 조치가 완료되는 데 약 30초가 걸립니다.
- Hue에서 쿼리에 대해 응답 시간이 평소보다 길어지는 경우 해당 쿼리를 다시 실행해 봅니다.

Hue용 고급 구성

이 단원에는 다음 주제가 포함되어 있습니다.

주제

LDAP 사용자를 위해 Hue 구성

LDAP과의 통합을 통해 사용자는 LDAP 디렉터리에 저장된 기존 자격 증명을 사용하여 Hue에 로그인할 수 있습니다. Hue를 LDAP과 통합할 때 Hue의 사용자 정보를 독립적으로 관리할 필요가 없습니다. 아래 정보에서는 Microsoft Active Directory와 Hue의 통합을 보여 주지만 구성 옵션은 모든 LDAP 디렉터리와 유사합니다.

LDAP 인증은 먼저 서버에 바인드하여 연결을 설정합니다. 그런 다음 설정된 연결은 이후의 모든 쿼리에 사용되어 LDAP 사용자 정보를 검색합니다. Active Directory 서버가 익명 연결을 허용하지 않으면 바인드 고유 이름과 암호를 사용하여 연결을 설정해야 합니다. 바인드 고유 이름(또는 DN)은 `bind_dn` 구성 설정에 의해 정의됩니다. 바인드 암호는 `bind_password` 구성 설정에 의해 정의됩니다. Hue에는 LDAP 요청을 바인딩하는 두 가지 방법인 검색 바인딩과 직접 바인딩이 있습니다. Amazon EMR과 함께 Hue를 사용하기 위해 선호되는 방법은 검색 바인드입니다.

Active Directory에서 검색 바인드를 사용하면 Hue는 `user_name_attr config`에 정의된 사용자 이름 속성을 사용하여 기본 고유 이름(또는 DN)에서 검색해야 하는 속성을 찾습니다. 검색 바인드는 전체 DN이 Hue 사용자에게 알려지지 않은 경우에 유용합니다.

예를 들어, 공통 이름(또는 CN)을 사용하도록 `user_name_attr config`을 설정했을 수 있습니다. 이 경우 Active Directory 서버는 로그인 중에 제공된 Hue 사용자 이름을 사용하여 디렉터리 트리에서 일치하는 공통 이름(기본 고유 이름에서 시작)을 검색합니다. Hue 사용자의 공통 이름이 발견되면 사용자의 고유 이름이 서버에 의해 반환됩니다. 그런 다음 Hue는 바인드 작업을 수행하여 사용자를 인증하는 데 사용되는 고유 이름을 구성합니다.

> **Note**
>
> 검색 바인드는 기본 고유 이름에서 시작하는 모든 디렉터리 하위 트리에서 사용자 이름을 검색합니다. Hue LDAP 구성에 지정된 기본 고유 이름은 사용자 이름과 가장 가까운 상위여야 하며 그렇지 않으면 LDAP 인증 성능이 저하될 수 있습니다.

Active Directory에서 직접 바인딩을 사용하는 경우 정확한 `nt_domain` 또는 `ldap_username_pattern`를 사용하여 인증해야 합니다. 직접 바인딩을 사용하는 경우 `nt_domain` 구성 설정으로 정의된 nt 도메인 속성이 정의되어 있으면 `<login username>@nt_domain` 형식의 양식을 사용하여 사용자 고유 이름 템플릿이 만들어집니다. 이 템플릿을 사용하여 기본 고유 이름에서 시작하는 모든 디렉터리 하위 트리를 검색합니다. nt 도메인이 구성되지 않은 경우, Hue는 사용자의 정확한 고유 이름 패턴(`ldap_username_pattern` 구성 설정에 의해 정의됨)을 검색합니다. 이 경우 서버는 기본 고유 이름에서 시작하는 모든 디렉터리 하위 트리에서 일치하는 `ldap_username_pattern` 값을 검색합니다.

AWS CLI를 사용하여 Hue용 LDAP 속성이 포함된 클러스터를 시작하는 방법

- `hue-ini`에 LDAP 속성을 지정하려면 Hue가 설치된 클러스터를 생성하고 LDAP용 구성 속성이 포함된 json 파일을 참조합니다. 예제 명령은 다음과 같습니다. 이 명령에서는 Amazon S3에 저장된 `myConfig.json` 구성 파일을 참조합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Hue Name=Spark
 Name=Hive \
--instance-type m5.xlarge --instance-count 3 --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json.
```

다음은 `myConfig.json`의 콘텐츠 예제입니다.

```
[
    {
        "Classification": "hue-ini",
        "Properties": {},
        "Configurations": [
            {
                "Classification": "desktop",
                "Properties": {},
                "Configurations": [
                    {
                        "Classification": "ldap",
```

```
                            "Properties": {},
                            "Configurations": [
                                {
                                    "Classification": "ldap_servers",
                                    "Properties": {},
                                    "Configurations": [
                                        {
                                            "Classification": "yourcompany",
                                            "Properties": {
                                                "base_dn":
 "DC=yourcompany,DC=hue,DC=com",
                                                "ldap_url": "ldap://ldapurl",
                                                "search_bind_authentication": "true",
                                                "bind_dn":
 "CN=hue,CN=users,DC=yourcompany,DC=hue,DC=com",
                                                "bind_password": "password"
                                            },
                                            "Configurations": []
                                        }
                                    ]
                                }
                            ]
                        },
                        {
                            "Classification": "auth",
                            "Properties": {
                             "backend": "desktop.auth.backend.LdapBackend"
                                 }
                        }
                    ]
                }

            ]
        }
]
```

> **Note**
>
> Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔,

AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

Hue에서 LDAP 설정을 보는 방법

1. 활성화된 VPN 연결 또는 Amazon EMR 클러스터의 프라이머리 노드에 대한 SSH 터널이 있는지 확인합니다. 그런 다음 브라우저에서 *master-public-dns*:8888을 입력하여 Hue 웹 인터페이스를 여십시오.
2. Hue 관리자 자격 증명을 사용하여 로그인합니다. 유용한 정보: 창이 열리면 Got it, prof!를 클릭하여 창을 닫습니다.
3. 도구 모음에서 Hue 아이콘을 클릭합니다.
4. Hue 정보 페이지에서 구성을 선택합니다.
5. Configuration Sections and Variables(구성 섹션 및 변수) 섹션에서 Desktop(데스크톱)을 클릭합니다.
6. ldap 섹션으로 스크롤하여 설정을 확인합니다.

Hue 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Hue의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Hue 버전 정보

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-7.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
		timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-7.6.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.5.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.4.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-7.3.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.2.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.2	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-7.1.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-7.0.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.15.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.14.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.13.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.12.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.10.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.10.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.9.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.9.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.8.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.8.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.7.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.36.1	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.36.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.6.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.35.0	4.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.3.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.3.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.2.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.2.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.1.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.1.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-6.0.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-6.0.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.34.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.33.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.33.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.32.1	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.32.0	4.8.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.31.1	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.31.0	4.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.30.2	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.30.1	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server
emr-5.30.0	4.6.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mariadb-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.29.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.28.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.28.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.27.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.27.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.26.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.25.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.24.1	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.24.0	4.4.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.23.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.23.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.22.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.21.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.21.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.20.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.20.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.19.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.19.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.18.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.18.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.17.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.17.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.17.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.16.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.16.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.15.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.15.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.14.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.14.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.13.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.13.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.12.3	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.2	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.12.1	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.12.0	4.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.4	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.3	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.11.2	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.11.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.10.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.10.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.9.1	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.9.0	4.0.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.8.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.8.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.7.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.7.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.6.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server
emr-5.6.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.5.4	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.3	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.2	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.5.1	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.5.0	3.12.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.4.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.4.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.2	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.3.1	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.3.0	3.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.2.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.2.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.1.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.1.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-5.0.3	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.2	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-5.0.1	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-5.0.0	3.10.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.9.6	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.9.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.9.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.9.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.5	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.8.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.8.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.8.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.4	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.7.3	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.2	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-client, oozie-server
emr-4.7.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.7.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.6.1	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.6.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.5.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.4.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.3.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server
emr-4.2.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Amazon EMR 릴리스 레이블	Hue 버전	Hue와 함께 설치된 구성 요소
emr-4.1.0	3.7.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hue-server, mysql-server, oozie-server

Iceberg

Apache Iceberg는 Amazon Simple Storage Service(S3)의 대형 데이터 세트를 위한 오픈 테이블 형식입니다. 이 테이블 형식은 대형 테이블, 원자성 커밋, 동시 쓰기, SQL 호환 테이블 진화 등에서 빠른 쿼리 성능을 제공합니다. Amazon EMR 6.5.0부터 Iceberg 테이블 형식으로 Amazon EMR 클러스터에서 Apache Spark 3을 사용할 수 있습니다.

다음 테이블에는 Amazon EMR이 Iceberg를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Iceberg의 버전이 나열되어 있습니다.

이 릴리스에서 Iceberg와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Iceberg 버전 정보

Amazon EMR 릴리스 레이블	Iceberg 버전	Iceberg와 함께 설치된 구성 요소
emr-7.9.0	Iceberg 1.7.1-amzn-2	Not available.

다음 테이블에는 Amazon EMR이 Iceberg를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Iceberg의 버전이 나열되어 있습니다.

이 릴리스에서 Iceberg와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Icebert 버전 정보

Amazon EMR 릴리스 레이블	Iceberg 버전	Iceberg와 함께 설치된 구성 요소
emr-6.15.0	Iceberg 1.4.0-amzn-0	Not available.

주제

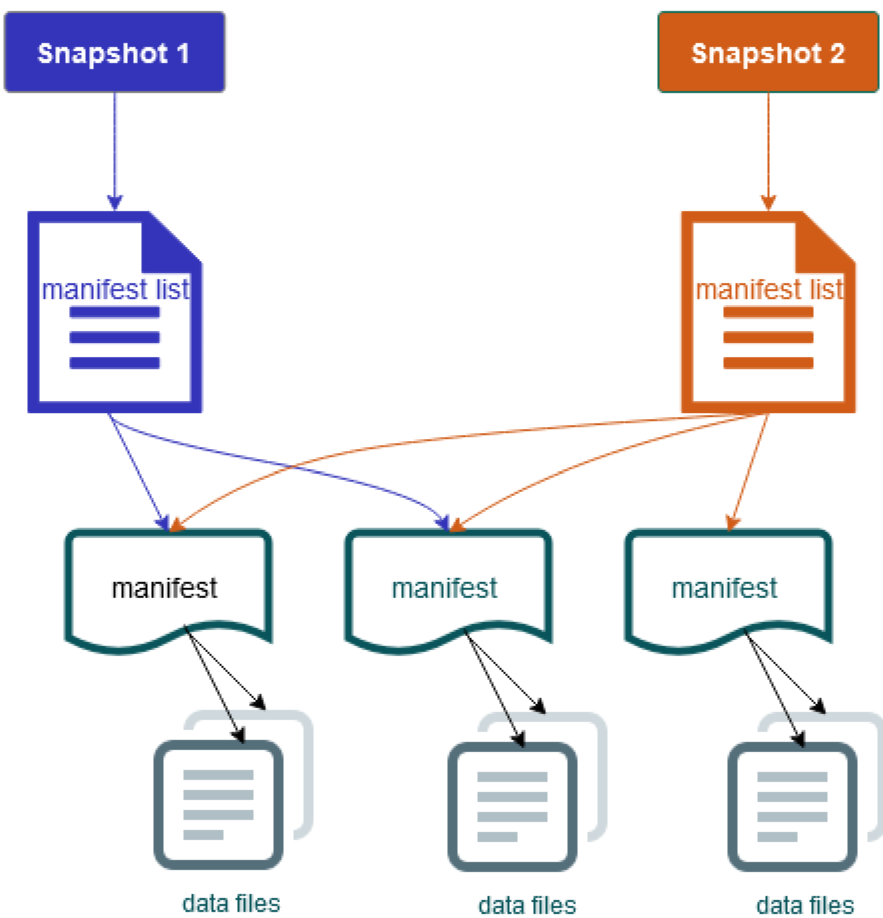
- Iceberg 작동 방식
- Iceberg와 함께 클러스터 사용
- Iceberg 릴리스 기록

Iceberg 작동 방식

Iceberg는 디렉터리가 아닌 테이블에서 개별 데이터 파일을 추적합니다. 이렇게 하면 작성자가 데이터 파일을 제자리에 생성할 수 있습니다(파일을 이동하거나 변경하지 않음). 또한 작성자는 명시적 커밋을 통해서만 테이블에 파일을 추가할 수 있습니다. 테이블 상태는 메타데이터 파일에서 유지 관리됩니다. 테이블 상태가 변경되면 항상 메타데이터를 자동으로 대체하는 새 메타데이터 파일이 생성됩니다. 테이블 메타데이터 파일은 테이블 스키마, 파티셔닝 구성 및 기타 속성을 추적합니다.

또한 테이블 콘텐츠의 스냅샷도 포함합니다. 각 스냅샷은 특정 시점에 테이블에 있는 전체 데이터 파일 세트입니다. 스냅샷은 메타데이터 파일에 나열되지만 스냅샷의 파일은 별도의 매니페스트 파일에 저장됩니다. 한 테이블 메타데이터 파일에서 다음 테이블 메타데이터 파일로의 원자적 전환은 스냅샷 격리를 제공합니다. 독자는 테이블 메타데이터를 로드했을 때 최신 상태였던 스냅샷을 사용합니다. 독자는 새로 고침을 통해 새 메타데이터 위치를 선택할 때까지 변경 사항의 영향을 받지 않습니다. 스냅샷의 데이터 파일은 테이블의 각 데이터 파일, 파티션 데이터 및 지표에 대한 행을 포함하는 하나 이상의 매니페스트 파일에 저장됩니다. 스냅샷은 매니페스트에 있는 모든 파일을 통합한 것입니다. 자주 변경되지 않는 메타데이터를 다시 작성하지 않도록 스냅샷 간에 매니페스트 파일을 공유할 수도 있습니다.

Iceberg 스냅샷 다이어그램



Iceberg는 다음 기능을 제공합니다.

- Amazon S3 데이터 레이크의 ACID 트랜잭션 및 시간 이동을 지원합니다.
- 커밋 재시도를 통해 낙관적 동시성의 성능 이점을 활용합니다.
- 파일 수준 충돌 해결은 높은 동시성을 제공합니다.
- 메타데이터의 열당 최소 및 최대 통계를 사용하면 파일을 건너뛸 수 있어서 선택적 쿼리의 성능이 향상됩니다.
- 파티션 스키마에 대한 업데이트를 활성화하는 파티션 진화를 통해 테이블을 유연한 파티션 레이아웃으로 구성할 수 있습니다. 그러면 실제 디렉터리에 의존하지 않고도 쿼리와 데이터 볼륨을 변경할 수 있습니다.
- 스키마 진화 및 적용을 지원합니다.
- Iceberg 테이블은 멱등성 싱크 및 재생 가능한 소스 역할을 합니다. 이를 통해 정확히 한 번의 파이프라인으로 스트리밍 및 배치 지원이 가능합니다. 멱등성 싱크는 이전에 성공했던 쓰기 작업을 추적합니다. 따라서 싱크는 장애 발생 시 데이터를 다시 요청하고, 여러 번 전송된 경우 데이터를 삭제할 수 있습니다.
- 테이블 진화, 작업 기록, 각 커밋에 대한 통계를 포함한 기록과 계보를 확인합니다.
- 원하는 데이터 형식(Parquet, ORC, Avro) 및 분석 엔진(Spark, Trino, PrestoDB, Flink, Hive)을 통해 기존 데이터 세트에서 마이그레이션합니다.

Iceberg와 함께 클러스터 사용

이 섹션에는 Spark, Trino, Flink, Hive와 함께 Iceberg를 사용하는 방법에 대한 정보가 포함되어 있습니다.

Spark와 함께 Iceberg 클러스터 사용

Amazon EMR 버전 6.5.0부터 부트스트랩 작업을 포함할 필요 없이 Spark 클러스터에서 Iceberg를 사용할 수 있습니다. Amazon EMR 버전 6.4.0 이하 버전의 경우 부트스트랩 작업을 사용하여 필요한 모든 종속 항목을 사전 설치할 수 있습니다.

이 자습서에서는 AWS CLI 를 사용하여 Amazon EMR Spark 클러스터에서 Iceberg로 작업합니다. 콘솔을 사용하여 Iceberg가 설치된 클러스터를 생성하려면 Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue의 단계를 수행합니다.

Iceberg 클러스터 생성

AWS Management Console AWS CLI 또는 Amazon EMR API를 사용하여 Iceberg가 설치된 클러스터를 생성할 수 있습니다. 이 자습서에서는 AWS CLI 를 사용하여 Amazon EMR 클러스터에서 Iceberg로 작업합니다. 콘솔을 사용하여 Iceberg가 설치된 클러스터를 생성하려면 Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue의 단계를 수행합니다.

와 함께 Amazon EMR에서 Iceberg를 사용하려면 AWS CLI먼저 다음 단계를 사용하여 클러스터를 생성합니다. 를 사용하여 Iceberg 분류를 지정하는 방법에 대한 자세한 내용은 클러스터를 생성할 AWS CLI 때를 사용하여 구성 제공 또는 섹션을 AWS CLI참조하세요클러스터를 생성할 때 Java SDK를 사용하여 구성 제공.

1. 다음 콘텐츠가 포함된 `configurations.json` 파일을 생성합니다.

```
[{
    "Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

2. 그리고 다음 구성을 사용하여 클러스터를 생성합니다. 예제 Amazon S3 버킷 경로와 서브넷 ID를 사용자 정보로 교체합니다.

```
aws emr create-cluster --release-label emr-6.5.0 \
--applications Name=Spark \
--configurations file://configurations.json \
--region us-east-1 \
--name My_Spark_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef0
```

또는 Spark 애플리케이션을 포함하는 Amazon EMR 클러스터를 생성하고 `/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar` 파일을 Spark 작업에 JAR 종속 항목으로 포함할 수 있습니다. 자세한 내용은 애플리케이션 제출을 참조하세요.

jar를 Spark 작업에 종속 항목으로 포함하려면 Spark 애플리케이션에 다음 구성 속성을 추가합니다.

```
--conf "spark.jars=/usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar"
```

Spark 작업 종속성에 대한 자세한 내용은 Apache Spark 설명서 Running Spark on Kubernetes에서 Dependency Management를 참조하세요.

Iceberg용 Spark 세션 초기화

다음 예제에서는 대화식 Spark 셸을 시작하거나 Spark 제출을 사용하거나 Amazon EMR에서 Iceberg를 작업하기 위해 Amazon EMR Notebooks를 사용하는 방법을 보여줍니다.

spark-shell

1. SSH를 사용하여 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. Spark 셸을 시작하려면 다음 명령을 입력합니다. PySpark 셸을 사용하려면 `spark-shell`을 `pyspark`로 바꿉니다.

```
spark-shell \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. SSH를 사용하여 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. Iceberg용 Spark 세션을 시작하려면 다음 명령을 입력합니다.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

EMR Studio 노트북을 사용하여 Spark 세션을 초기화하려면 다음 예제와 같이 Amazon EMR Notebooks에서 `%%configure` 매직 명령을 사용하여 Spark 세션을 구성합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR Notebooks 매직 사용을 참조하세요.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
}
```

CLI

CLI를 사용하여 Spark 클러스터를 초기화하고 모든 Spark Iceberg 세션 기본 구성을 설정하려면 다음 샘플을 실행합니다. AWS CLI 및 Amazon EMR API를 사용하여 구성 분류를 지정하는 방법에 대한 자세한 내용은 애플리케이션 구성을 참조하세요.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

Iceberg 테이블에 쓰기

다음 예제에서는 Spark DataFrame을 생성하고 Iceberg 데이터 세트로 쓰는 방법을 보여줍니다. 이 예제에서는 SSH를 기본 hadoop 사용자로 사용하여 프라이머리 노드에 연결되어 있는 동안 Spark 셸을 통해 데이터 세트를 작업하는 방법을 보여줍니다.

> **Note**
>
> 코드 샘플을 Spark 셸에 붙여넣으려면 프롬프트에 `:paste`를 입력하고 예제를 붙여넣은 다음 `CTRL+D`를 누릅니다.

PySpark

Spark에는 Python 기반 셸(`pyspark`)이 포함되어 있으며, 이 셸을 사용하여 Python에서 작성된 Spark 프로그램을 시제품화할 수 있습니다. 프라이머리 노드에서 `pyspark`를 간접 호출합니다.

```
## Create a DataFrame.
data = spark.createDataFrame([
 ("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
 ("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
 ("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
 ("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
],["id", "creation_date", "last_update_time"])

## Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

Scala

```
import org.apache.spark.sql.SaveMode
import org.apache.spark.sql.functions._

// Create a DataFrame.
val data = Seq(
```

```
("100", "2015-01-01", "2015-01-01T13:51:39.340396Z"),
("101", "2015-01-01", "2015-01-01T12:14:58.597216Z"),
("102", "2015-01-01", "2015-01-01T13:51:40.417052Z"),
("103", "2015-01-01", "2015-01-01T13:51:40.519832Z")
).toDF("id", "creation_date", "last_update_time")

// Write a DataFrame as a Iceberg dataset to the Amazon S3 location.
spark.sql("""CREATE TABLE IF NOT EXISTS dev.db.iceberg_table (id string,
creation_date string,
last_update_time string)
USING iceberg
location 's3://amzn-s3-demo-bucket/example-prefix/db/iceberg_table'""")

data.writeTo("dev.db.iceberg_table").append()
```

Iceberg 테이블에서 읽기

PySpark

```
df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Scala

```
val df = spark.read.format("iceberg").load("dev.db.iceberg_table")
df.show()
```

Spark SQL

```
SELECT * from dev.db.iceberg_table LIMIT 10
```

Spark Iceberg에서 AWS Glue 데이터 카탈로그 사용

Spark Iceberg에서 AWS Glue 데이터 카탈로그에 연결할 수 있습니다. 이 섹션에서는 서로 다른 연결 명령을 보여줍니다.

기본 리전의 기본 AWS Glue 카탈로그에 연결

이 샘플은 Glue 카탈로그 유형을 사용하여 연결하는 방법을 보여줍니다. 카탈로그 ID를 지정하지 않으면 기본값을 사용합니다.

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

특정 카탈로그 ID를 사용하여 AWS Glue 카탈로그에 연결

이 샘플은 카탈로그 ID를 사용하여 연결하는 방법을 보여줍니다.

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=AWS Glue catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

이 명령은 다른 계정의 AWS Glue 카탈로그, RMS 카탈로그 또는 페더레이션 카탈로그에 연결하는 데 사용할 수 있습니다.

Spark Iceberg에서 Iceberg REST 카탈로그(IRC) 사용

다음 섹션에서는 카탈로그와의 Iceberg 통합을 구성하는 방법을 자세히 설명합니다.

AWS Glue Data Catalog IRC 엔드포인트에 연결

다음은 Iceberg REST를 사용하기 위한 샘플 `spark-submit` 명령을 보여줍니다.

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=glue catalog ID \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=glue endpoint URI/iceberg \
    --conf spark.sql.catalog.my_catalog.rest.sigv4-enabled=true \
    --conf spark.sql.catalog.my_catalog.rest.signing-name=glue \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

런타임 역할이 활성화된 클러스터에서 사용하려면 다음과 같은 추가 Spark 구성 설정이 필요합니다.

```
"spark.hadoop.fs.s3.credentialsResolverClass":
 "software.amazon.glue.GlueTableCredentialsResolver",
"spark.hadoop.catalog-impl": "org.apache.iceberg.aws.glue.GlueCatalog",
"spark.hadoop.glue.id": glue catalog ID
"spark.hadoop.glue.endpoint": "glue endpoint"
```

각 리전의 AWS Glue 엔드포인트 URL 목록은 AWS Glue 엔드포인트 및 할당량을 참조하세요.

임의의 IRC 엔드포인트에 연결

다음은 IRC 엔드포인트를 사용하기 위한 샘플 `spark-submit` 명령을 보여줍니다.

```
spark-submit \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=warehouse name \
    --conf spark.sql.catalog.my_catalog.type=rest \
    --conf spark.sql.catalog.my_catalog.uri=your rest endpoint \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Iceberg SparkCatalog를 사용할 때와 SparkSessionCatalog를 사용할 때의 구성 차이

Iceberg는 Spark Iceberg 카탈로그를 생성하는 두 가지 방법을 제공합니다. Spark 구성을 `SparkCatalog` 또는 로 설정할 수 있습니다`SparkSessionCatalog`.

Iceberg SparkCatalog 사용

다음은 SparkCatalog를 Spark Iceberg 카탈로그로 사용하기 위한 명령을 보여줍니다.

```
spark-shell \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.defaultCatalog=my_catalog
```

이 접근 방식에 대한 고려 사항:

- Iceberg 테이블에는 액세스할 수 있지만 다른 테이블은 액세스할 수 없습니다.

- 카탈로그 이름은 spark_catalog일 수 없습니다. Spark의 초기 카탈로그 이름입니다. 항상 Hive 메타스토어에 연결됩니다. 사용자가를 사용하여 덮어쓰지 않는 한 Spark의 기본 카탈로그입니다`spark.sql.defaultCatalog`.
- `spark.sql.defaultCatalog`를 카탈로그 이름으로 설정하여 이를 기본 카탈로그로 만들 수 있습니다.

Iceberg SparkSessionCatalog 사용

다음은 SparkSessionCatalog를 Spark Iceberg 카탈로그로 사용하기 위한 명령을 보여줍니다.

```
spark-shell \
    --conf spark.sql.catalog.spark_catalog=org.apache.iceberg.spark.SparkSessionCatalog
 \
    --conf spark.sql.catalog.spark_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
    --conf spark.sql.catalog.spark_catalog.type=glue
```

이 접근 방식에 대한 고려 사항:

- 테이블을 Iceberg 테이블로 찾을 수 없는 경우 Spark는 해당 테이블이 Hive 메타스토어의 테이블인지 확인합니다. 자세한 내용은 AWS Glue 데이터 카탈로그를 Hive용 카탈로그로 사용을 참조하세요.
- 카탈로그 이름은 spark_catalog여야 합니다.

Iceberg Spark 확장 사용

Iceberg는 사용자가 Spark 확장 구성를 통해 설정할 수 `org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions` 있는 Spark 확장을 제공합니다`spark.sql.extensions`. 확장을 통해 행 수준 DELETE, UPDATE 및 MERGE와 같은 주요 Iceberg 기능, 압축, 스냅샷 만료, 분기 및 태그 지정 등과 같은 Iceberg별 Spark 데이터 정의 언어 설명 및 절차를 사용할 수 있습니다. 자세한 내용은 다음을 참조하세요.

- Iceberg Spark 쓰기 확장: Spark 쓰기
- Iceberg Spark DDL 확장: ALTER TABLE SQL 확장
- Iceberg Spark 프로시저 확장: Spark 프로시저

Spark에서 Iceberg 사용 시 고려 사항

- Amazon EMR 6.5.0에서는 기본적으로 Amazon EMR on EKS에서 실행되는 Iceberg를 지원하지 않습니다. Amazon EMR 6.5.0 사용자 지정 이미지를 사용할 수 있으므로 Amazon EMR on EKS에서 Iceberg 테이블을 생성하기 위해 `--jars local:///usr/share/aws/iceberg/lib/iceberg-spark3-runtime.jar`을 `spark-submit` 파라미터로 전달할 수 있습니다. 자세한 내용은 Amazon EMR on EKS 개발 안내서에서 [사용자 지정 이미지를 사용하여 Amazon EMR에서 Spark 워크로드 제출]을 참조하세요. 지원 에 지원을 문의할 수도 있습니다. Amazon EMR 6.6.0부터 Amazon EMR on EKS에서 Iceberg가 지원됩니다.
- AWS Glue를 Iceberg의 카탈로그로 사용하는 경우 테이블을 생성하는 데이터베이스가 AWS Glue에 있는지 확인합니다. AWS Lake Formation 와 같은 서비스를 사용하고 있고 카탈로그를 로드할 수 없는 경우 명령을 실행하기 위해 서비스에 대한 적절한 액세스 권한이 있는지 확인합니다.
- 에 설명된 대로 Iceberg SparkSessionCatalog를 사용하는 경우 Spark Iceberg AWS Glue 데이터 카탈로그 설정을 구성하는 것 외에도 Glue 데이터 카탈로그를 Apache Hive 메타스토어로 구성에 설명된 구성 단계를 따라야 Iceberg SparkCatalog를 사용할 때와 SparkSessionCatalog를 사용할 때의 구성 차이합니다. AWS

Trino와 함께 Iceberg 클러스터 사용

Amazon EMR 버전 6.6.0부터 Iceberg를 Trino 클러스터와 함께 사용할 수 있습니다.

이 자습서에서는 AWS CLI 를 사용하여 Amazon EMR Trino 클러스터에서 Iceberg로 작업합니다. 콘솔을 사용하여 Iceberg가 설치된 클러스터를 생성하려면 Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue의 단계를 수행합니다.

Iceberg 클러스터 생성

와 함께 Amazon EMR에서 Iceberg를 사용하려면 AWS CLI먼저 다음 단계를 사용하여 클러스터를 생성합니다. 를 사용하여 Iceberg 분류를 지정하는 방법에 대한 자세한 내용은 클러스터를 생성할 AWS CLI 때를 사용하여 구성 제공 또는 섹션을 AWS CLI참조하세요클러스터를 생성할 때 Java SDK를 사용하여 구성 제공.

1. `iceberg.properties` 파일을 생성하고 선택한 카탈로그의 값을 설정합니다. 예를 들어 Hive 메타스토어를 카탈로그로 사용하려면 파일에 다음과 같은 콘텐츠가 있어야 합니다.

```
connector.name=iceberg
hive.metastore.uri=thrift://localhost:9083
```

AWS Glue 데이터 카탈로그를 스토어로 사용하려면 파일에 다음 콘텐츠가 있어야 합니다.

```
connector.name=iceberg
iceberg.catalog.type=glue
```

2. 다음 예제와 같이 `iceberg.properties`를 Amazon S3에서 `/etc/trino/conf/catalog/iceberg.properties`로 복사하는 부트스트랩 작업을 생성합니다. 부트스트랩 작업에 대한 자세한 내용은 부트스트랩 작업을 생성하여 추가 소프트웨어 설치를 참조하세요.

```
set -ex
sudo aws s3 cp s3://amzn-s3-demo-bucket/iceberg.properties /etc/trino/conf/catalog/
iceberg.properties
```

3. 다음과 같은 구성으로 클러스터를 생성해 예제 부트스트랩 작업 스크립트 경로 및 키 이름을 사용자 정보로 대체합니다.

```
aws emr create-cluster --release-label emr-6.7.0 \
--applications Name=Trino \
--region us-east-1 \
--name My_Trino_Iceberg_Cluster \
--bootstrap-actions '[{"Path":"s3://amzn-s3-demo-bucket","Name":"Add
 iceberg.properties"}]' \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=c3.4xlarge
 InstanceGroupType=CORE,InstanceCount=3,InstanceType=c3.4xlarge \
--use-default-roles \
--ec2-attributes KeyName=<key-name>
```

Iceberg용 Trino 세션 초기화

Trino 세션을 초기화하려면 다음 명령을 실행합니다.

```
trino-cli --catalog iceberg
```

Iceberg 테이블에 쓰기

다음 SQL 명령을 사용하여 생성 후 테이블에 작성합니다.

```
trino> SHOW SCHEMAS;
trino> CREATE TABLE default.iceberg_table (
```

```
            id int,
            data varchar,
            category varchar)
        WITH (
            format = 'PARQUET',
            partitioning = ARRAY['category', 'bucket(id, 16)'],
            location = 's3://amzn-s3-demo-bucket/<prefix>')

trino> INSERT INTO default.iceberg_table VALUES (1,'a','c1'), (2,'b','c2'),
 (3,'c','c3');
```

Iceberg 테이블에서 읽기

Iceberg 테이블에서 읽으려면 다음 명령을 실행합니다.

```
trino> SELECT * from default.iceberg_table;
```

Trino에서 Iceberg 사용 시 고려 사항

- Amazon EMR 6.5는 기본적으로 Iceberg에 대한 Trino Iceberg 카탈로그 지원을 제공하지 않습니다. Trino에는 Iceberg v0.11이 필요하므로 Spark 클러스터와 별도로 Trino용 Amazon EMR 클러스터를 시작하고 해당 클러스터에 Iceberg v0.11을 포함하는 것이 좋습니다.
- AWS Glue를 Iceberg의 카탈로그로 사용하는 경우 테이블을 생성하는 데이터베이스가 AWS Glue에 있는지 확인합니다. AWS Lake Formation 와 같은 서비스를 사용하고 있고 카탈로그를 로드할 수 없는 경우 명령을 실행하기 위해 서비스에 대한 적절한 액세스 권한이 있는지 확인합니다.
- Iceberg Glue 통합은 Redshift Managed Storage 카탈로그에서 작동하지 않습니다.

Flink와 함께 Iceberg 클러스터 사용

Amazon EMR 버전 6.9.0부터 오픈 소스 Iceberg Flink 통합을 사용할 때 필수 설정 단계 없이도 Flink 클러스터에서 Iceberg를 사용할 수 있습니다.

Iceberg 클러스터 생성

AWS Management Console, AWS CLI또는 Amazon EMR API를 사용하여 Iceberg가 설치된 클러스터를 생성할 수 있습니다. 이 자습서에서는 AWS CLI 를 사용하여 Amazon EMR 클러스터에서 Iceberg로 작업합니다. 콘솔을 사용하여 Iceberg가 설치된 클러스터를 생성하려면 [Build an Apache Iceberg data lake using Amazon Athena, Amazon EMR, and AWS Glue]의 단계를 수행합니다.

와 함께 Amazon EMR에서 Iceberg를 사용하려면 AWS CLI먼저 다음 단계를 사용하여 클러스터를 생성합니다. 를 사용하여 Iceberg 분류를 지정하는 방법에 대한 자세한 내용은 클러스터를 생성할 AWS CLI 때를 사용하여 구성 제공 또는 섹션을 AWS CLI참조하세요클러스터를 생성할 때 Java SDK를 사용하여 구성 제공. 다음 콘텐츠로 `configurations.json`이라는 파일을 생성합니다.

```
[{
"Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

그리고 다음과 같은 구성으로 클러스터를 생성하고, 예제 Amazon S3 버킷 경로와 서브넷 ID를 사용자 값으로 대체합니다.

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Flink \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_flink_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

또는 Flink 애플리케이션을 포함하는 Amazon EMR 6.9.0 클러스터를 생성하고 `/usr/share/aws/iceberg/lib/iceberg-flink-runtime.jar` 파일을 Flink 작업에 JAR 종속성으로 포함할 수 있습니다.

Flink SQL 클라이언트 사용

SQL 클라이언트 스크립트는 `/usr/lib/flink/bin`에 있습니다. 다음 명령으로 스크립트를 실행할 수 있습니다.

```
flink-yarn-session -d # starting the Flink YARN Session in detached mode
./sql-client.sh
```

그러면 Flink SQL 셸이 시작됩니다.

Flink 예제

Iceberg 테이블 생성

Flink SQL

```
CREATE CATALOG glue_catalog WITH (
   'type'='iceberg',
   'warehouse'='<WAREHOUSE>',
   'catalog-type'='glue'
 );

USE CATALOG  glue_catalog;

CREATE DATABASE IF NOT EXISTS <DB>;

USE <DB>;

CREATE TABLE IF NOT EXISTS `glue_catalog`.`<DB>`.`sample` (id int, data string);
```

테이블 API

```
EnvironmentSettings settings =
                EnvironmentSettings.newInstance().inBatchMode().build();

TableEnvironment tEnv = TableEnvironment.create(settings);

String warehouse = "<WAREHOUSE>";
String db = "<DB>";

tEnv.executeSql(
                "CREATE CATALOG glue_catalog WITH (\n"
                        + "   'type'='iceberg',\n"
                        + "   'warehouse'='"
                        + warehouse
                        + "',\n"
                        + "   'catalog-impl'='org.apache.iceberg.aws.glue.GlueCatalog',
\n"
                        + "   'io-impl'='org.apache.iceberg.aws.s3.S3FileIO'\n"
                        + " );");

tEnv.executeSql("USE CATALOG  glue_catalog;");
tEnv.executeSql("CREATE DATABASE IF NOT EXISTS " + db + ";");
```

```
tEnv.executeSql("USE " + db + ";");
tEnv.executeSql(
        "CREATE TABLE `glue_catalog`.`" + db + "`.`sample` (id bigint, data string);");
```

Iceberg 테이블에 쓰기

Flink SQL

```
INSERT INTO `glue_catalog`.`<DB>`.`sample` values (1, 'a'),(2,'b'),(3,'c');
```

테이블 API

```
tEnv.executeSql(
        "INSERT INTO `glue_catalog`.`"
                + db
                + "`.`sample` values (1, 'a'),(2,'b'),(3,'c');");
```

데이터스트림 API

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

GenericRowData rowData1 = new GenericRowData(2);
rowData1.setField(0, 1L);
rowData1.setField(1, StringData.fromString("a"));

DataStream<RowData> input = env.fromElements(rowData1);

Map<String, String> props = new HashMap<();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
```

```
                props,
                new Configuration(),
                "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));

DataStreamSink<Void> dataStreamSink =
        FlinkSink.forRowData(input).tableLoader(tableLoader).append();

env.execute("Datastream Write");
```

Iceberg 테이블에서 읽기

Flink SQL

```
SELECT * FROM `glue_catalog`.`<DB>`.`sample`;
```

테이블 API

```
Table result = tEnv.sqlQuery("select * from `glue_catalog`.`" + db + "`.`sample`;");
```

데이터스트림 API

```
final StreamExecutionEnvironment env =
 StreamExecutionEnvironment.getExecutionEnvironment();

final StreamTableEnvironment tableEnv = StreamTableEnvironment.create(env);

String db = "<DB Name>";

String warehouse = "<Warehouse Path>";

Map<String, String> props = new HashMap<>();
props.put("type", "iceberg");
props.put("warehouse", warehouse);
props.put("io-impl", "org.apache.iceberg.aws.s3.S3FileIO");

CatalogLoader glueCatlogLoader =
        CatalogLoader.custom(
                "glue",
                props,
                new Configuration(),
```

```
                 "org.apache.iceberg.aws.glue.GlueCatalog");

TableLoader tableLoader =
        TableLoader.fromCatalog(glueCatlogLoader, TableIdentifier.of(db, "sample"));

DataStream<RowData> batch =

 FlinkSource.forRowData().env(env).tableLoader(tableLoader).streaming(false).build();

batch.print().name("print-sink");
```

Hive 카탈로그 사용

Hive 메타스토어 및 Glue 카탈로그를 사용하여 Flink 구성에 설명된 대로 Flink 및 Hive 종속성이 해결되었는지 확인합니다.

Flink 작업 실행

Flink에 작업을 제출하는 한 가지 방법은 작업별 Flink YARN 세션을 사용하는 것입니다. 다음 명령을 사용하여 시작할 수 있습니다.

```
sudo flink run -m yarn-cluster -p 4 -yjm 1024m -ytm 4096m $JAR_FILE_NAME
```

Flink에서 Iceberg 사용 시 고려 사항

- AWS Glue를 Iceberg의 카탈로그로 사용하는 경우 테이블을 생성하는 데이터베이스가 AWS Glue에 있는지 확인합니다. AWS Lake Formation 와 같은 서비스를 사용하고 있고 카탈로그를 로드할 수 없는 경우 명령을 실행하기 위해 서비스에 대한 적절한 액세스 권한이 있는지 확인합니다.
- Iceberg Glue 통합은 Redshift Managed Storage 카탈로그에서 작동하지 않습니다.

Iceberg 클러스터를 Hive와 함께 사용

Amazon EMR 릴리스 6.9.0 이상에서 오픈 소스 Iceberg Hive 통합에 필요한 설정 단계를 수행하지 않고도 Hive 클러스터에서 Iceberg를 사용할 수 있습니다. Amazon EMR 버전 6.8.0 이하 버전의 경우 부트스트랩 작업을 사용하여 Iceberg 지원을 위해 Hive를 구성하도록 `iceberg-hive-runtime` jar을 설치할 수 있습니다.

Amazon EMR 6.9.0에는Hive 3.1.3 integration with Iceberg 0.14.1의 모든 기능이 포함되어 있으며, 런타임 시 지원되는 실행 엔진 자동 선택과 같은 Amazon EMR 추가 기능(Amazon EMR on EKS 6.9.0)도 포함되어 있습니다.

Iceberg 클러스터 생성

AWS Management Console AWS CLI 또는 Amazon EMR API를 사용하여 Iceberg가 설치된 클러스터를 생성할 수 있습니다. 이 자습서에서는 AWS CLI 를 사용하여 Amazon EMR 클러스터에서 Iceberg로 작업합니다. 콘솔을 사용하여 Iceberg가 설치된 클러스터를 생성하려면 Amazon Athena, Amazon EMR 및 AWS Glue를 사용하여 Iceberg 데이터 레이크 구축의 단계를 따릅니다.

와 함께 Amazon EMR에서 Iceberg를 사용하려면 AWS CLI먼저 아래 단계를 사용하여 클러스터를 생성합니다. AWS CLI 또는 Java SDK를 사용하여 Iceberg 분류를 지정하는 방법에 대한 자세한 내용은 클러스터를 생성할 AWS CLI 때를 사용하여 구성 제공 또는 섹션을 참조하세요클러스터를 생성할 때 Java SDK를 사용하여 구성 제공. 다음 콘텐츠가 포함된 `configurations.json`이라는 파일을 생성합니다.

```
[{
    "Classification":"iceberg-defaults",
    "Properties":{"iceberg.enabled":"true"}
}]
```

그리고 다음과 같은 구성으로 클러스터를 생성하고, 예제 Amazon S3 버킷 경로와 서브넷 ID를 사용자 정보로 대체합니다.

```
aws emr create-cluster --release-label emr-6.9.0 \
--applications Name=Hive \
--configurations file://iceberg_configurations.json \
--region us-east-1 \
--name My_hive_Iceberg_Cluster \
--log-uri s3://amzn-s3-demo-bucket/ \
--instance-type m5.xlarge \
--instance-count 2 \
--service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-1234567890abcdef
```

Hive Iceberg 클러스터는 다음과 같은 작업을 수행합니다.

- Iceberg Hive 런타임 jar을 Hive에 로드하고 Hive 엔진의 Iceberg 관련 구성을 활성화합니다.
- Amazon EMR Hive의 동적 실행 엔진 선택을 활성화하여 사용자가 Iceberg 호환성을 위해 지원되는 실행 엔진을 설정하지 못하도록 합니다.

> **Note**
>
> Hive Iceberg 클러스터는 현재 AWS Glue 데이터 카탈로그를 지원하지 않습니다. 기본 Iceberg 카탈로그는 `HiveCatalog`이며, 이는 Hive 환경에 맞게 구성된 메타스토어에 해당합니다. 카탈로그 관리에 대한 자세한 내용은 Apache Hive 설명서의 Using HCatalog를 참조하세요.

기능 지원

Amazon EMR 6.9.0은 Hive 3.1.3 및 Iceberg 0.14.1을 지원합니다. 이 기능은 Hive 3.1.2 및 3.1.3의 Iceberg 호환 기능으로 제한됩니다. 다음 명령이 지원됩니다.

- Amazon EMR 릴리스 6.9.0에서 6.12.x까지 Hive `auxlib` 디렉터리에 `libfb303` jar을 포함해야 합니다. 다음 명령을 사용하여 포함합니다.

```
sudo /usr/bin/ln -sf /usr/lib/hive/lib/libfb303-*.jar /usr/lib/hive/auxlib/
libfb303.jar
```

  Amazon EMR 릴리스 6.13 이상에서 `libfb303` jar은 Hive `auxlib` 디렉터리에 기호 링크로 자동 설정됩니다.

- 테이블 생성
  - 파티셔닝되지 않은 테이블 - 다음과 같이 스토리지 핸들러를 제공하여 Hive의 외부 테이블을 생성할 수 있습니다.

```
CREATE EXTERNAL TABLE x (i int) STORED BY
 'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
```

  - 파티셔닝된 테이블 - 다음과 같이 Hive에서 파티셔닝된 외부 테이블을 생성할 수 있습니다.

```
CREATE EXTERNAL TABLE x (i int) PARTITIONED BY (j int) STORED BY
 'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
```

> **Note**
>
> ORC, AVRO, PARQUET의 `STORED AS` 파일 형식은 Hive 3에서 지원되지 않습니다. 기본 옵션이자 유일한 옵션은 Parquet입니다.

- 테이블 삭제 - `DROP TABLE` 명령은 다음 예제와 같이 테이블을 삭제하는 데 사용됩니다.

```
DROP TABLE [IF EXISTS] table_name [PURGE];
```

- 테이블 읽기 - 다음 예제와 같이 SELECT 문을 사용하여 Hive에서 Iceberg 테이블을 읽을 수 있습니다. 지원되는 실행 엔진은 MR 및 Tez입니다.

  ```
  SELECT * FROM table_name
  ```

  Hive의 select 구문에 대한 자세한 내용은 LanguageManual Select를 참조하세요. Hive의 Iceberg 테이블에서 select 문에 대한 자세한 내용은 Apache Iceberg Select를 참조하세요.

- 테이블에 삽입 - HiveQL의 INSERT INTO 문은 Map Reduce 실행 엔진만 지원하는 Iceberg 테이블에서 작동합니다. Amazon EMR Hive는 런타임 시 Iceberg 테이블의 엔진을 선택하기 때문에 Amazon EMR 사용자는 실행 엔진을 명시적으로 설정할 필요가 없습니다.
  - 단일 테이블 삽입 - 예제:

    ```
    INSERT INTO table_name VALUES ('a', 1);
    INSERT INTO table_name SELECT...;
    ```

  - 다중 테이블 삽입 - 명령문에 비원자성 다중 테이블 삽입이 지원됩니다. 예시

    ```
    FROM source
     INSERT INTO table_1 SELECT a, b
     INSERT INTO table_2 SELECT c,d;
    ```

Amazon EMR 7.3.0부터 Iceberg를 사용하는 Hive는 AWS Glue 데이터 카탈로그를 메타스토어로 지원합니다. Glue 데이터 카탈로그를 AWS 메타스토어로 사용하려면 다음 속성을 설정합니다.

```
SET iceberg.catalog.<catalog_name>.catalog-
impl=org.apache.iceberg.aws.glue.GlueCatalog;
```

또는 다음 속성을 설정할 수도 있습니다.

```
SET iceberg.catalog.<catalog_name>.type=glue;
```

그리고 다음 예제를 사용하여 테이블을 생성할 수 있습니다.

```
CREATE EXTERNAL TABLE table_name (col1 type1, col2 type2,..)
ROW FORMAT SERDE 'org.apache.iceberg.mr.hive.HiveIcebergSerDe'
```

```
STORED BY 'org.apache.iceberg.mr.hive.HiveIcebergStorageHandler'
location '<location>'
TBLPROPERTIES ('table_type'='iceberg', 'iceberg.catalog'='<catalog_name>');
```

Hive에서 Iceberg 사용 시 고려 사항

- Iceberg는 다음과 같은 쿼리 유형을 지원합니다.
  - 테이블 생성
  - 테이블 삭제
  - 테이블에 삽입
  - 테이블 읽기
- MR(MapReduce) 실행 엔진만 데이터 조작 언어(DML) 작업에서 지원되며 MR은 Hive 3.1.3에서 더 이상 사용되지 않습니다.
- AWS Glue Data Catalog는 현재 Hive를 사용하는 Iceberg에서 지원되지 않습니다.
- 오류 처리 성능이 부족합니다. 구성이 잘못되었어도 쿼리에 삽입 작업은 성공적으로 완료될 수 있습니다. 그러나 메타데이터를 업데이트하지 않으면 데이터가 손실될 수 있습니다.
- Iceberg Glue 통합은 Redshift Managed Storage 카탈로그에서 작동하지 않습니다.

Iceberg 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Iceberg의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Iceberg 버전 정보

Amazon EMR 릴리스 레이블	Iceberg 버전	Iceberg와 함께 설치된 구성 요소
emr-7.9.0	1.7.1-amzn-2	Not available.
emr-7.8.0	1.7.1-amzn-1	Not available.

Amazon EMR 릴리스 레이블	Iceberg 버전	Iceberg와 함께 설치된 구성 요소
emr-7.7.0	1.7.1-amzn-0	Not available.
emr-7.6.0	1.6.1-amzn-2	Not available.
emr-7.5.0	1.6.1-amzn-1	Not available.
emr-7.4.0	1.6.1-amzn-0	Not available.
emr-7.3.0	1.5.2-amzn-0	Not available.
emr-7.2.0	1.5.0-amzn-0	Not available.
emr-7.1.0	1.4.3-amzn-0	Not available.
emr-7.0.0	1.4.2-amzn-0	Not available.
emr-6.15.0	1.4.0-amzn-0	Not available.
emr-6.14.0	1.3.1-amzn-0	Not available.
emr-6.13.0	1.3.0-amzn-1	Not available.
emr-6.12.0	1.3.0-amzn-0	Not available.
emr-6.11.1	1.2.0-amzn-0	Not available.
emr-6.11.0	1.2.0-amzn-0	Not available.
emr-6.10.1	1.1.0-amzn-0	Not available.
emr-6.10.0	1.1.0-amzn-0	Not available.
emr-6.9.1	0.14.1-amzn-0	Not available.
emr-6.9.0	0.14.1-amzn-0	Not available.
emr-6.8.1	0.14.0-amzn-0	Not available.
emr-6.8.0	0.14.0-amzn-0	Not available.

Amazon EMR 릴리스 레이블	Iceberg 버전	Iceberg와 함께 설치된 구성 요소
emr-6.7.0	0.13.1-amzn-0	Not available.
emr-6.6.0	0.13.1	Not available.
emr-6.5.0	0.12.0	Not available.

버전별 Iceberg 릴리스 정보

- Amazon EMR 6.9.0 - Iceberg 릴리스 정보

Amazon EMR 6.9.0 - Iceberg 릴리스 정보

Amazon EMR 6.9.0 - Iceberg 변경

유형	설명
Feature	Amazon EMR Flink와 Iceberg 통합.
Feature	Amazon EMR Hive와 Iceberg의 통합.
Feature	Amazon FSx for Lustre에서 Iceberg 메타데이터 파일 캐싱을 지원하여 쿼리 계획 시간을 개선합니다.
백포트	PR 5050: Flink 1.15: 인라인 삽입 SQL 주석의 쓰기 옵션을 지원합니다.
백포트	PR 5282: AWS: 새 데이터 파일 스트림을 열어 PUT 재시도 실패를 수정합니다.
백포트	PR 5318: Flink 1.15: FlinkSource 및 IcebergSource(FLIP-27) 사이의 격차를 해소하고 Flink SQL에서 FLIP-27 소스를 사용하기 위한 옵트인 구성을 추가했습니다.

유형	설명
백포트	PR 5344: Flink 1.14: FlinkSource 및 IcebergSource(FLIP-27) 사이의 격차를 해소하고 Flink SQL에서 FLIP-27 소스를 사용하기 위한 옵트인 구성을 추가했습니다.
백포트	PR 5393: Flink 1.14, 1.15: FLIP-27 소스 리더에서 Iceberg MetricContext를 Flink 지표로 변환하지 않습니다.
백포트	PR 5401: Flink 1.14, 1.15: FLIP-27 소스 리더 지표에 대한 PR 5393에서 IcebergSourceReader 그룹이 누락되었습니다.
백포트	PR 5679: Spark 3.2, 3.3: MergeRows 노드의 null 허용 여부 전파를 수정합니다.
백포트	PR 5860: Spark 3.3: 날짜로 파티셔닝된 테이블에서 RewriteManifestProcedure를 실행할 때 발생하는 쿼리 오류를 수정합니다.
백포트	PR 5880: Spark 3.3: 읽기 시 병합 프로젝션의 null 허용 여부를 수정합니다.
백포트	PR 5917: Spark 3.2: 읽기 시 병합 프로젝션의 null 허용 여부를 수정합니다.

Amazon EMR의 Jupyter Notebook

Jupyter Notebook은 실시간 코드, 수식, 가상화 및 설명 텍스트가 포함된 문서를 생성 및 공유하는 데 사용할 수 있는 오픈 소스 웹 애플리케이션입니다. Amazon EMR은 Jupyter Notebook으로 작업할 수 있는 세 가지 옵션을 제공합니다.

주제

EMR Studio

Amazon EMR Studio는 Amazon EMR 클러스터에서 실행되는 완전관리형 Jupyter Notebook을 위한 웹 기반 통합 개발 환경(IDE)입니다. 팀이 R, Python, Scala 및 PySpark로 작성된 애플리케이션을 개발, 시각화 및 디버깅할 수 있도록 EMR Studio를 설정할 수 있습니다.

Amazon EMR에서 Jupyter Notebook을 사용할 때는 EMR Studio를 사용하는 것이 좋습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR Studio를 참조하세요.

Jupyter Notebook에 기반한 Amazon EMR 노트북

EMR Notebooks는 Amazon EMR 콘솔에 내장된 Jupyter Notebook 환경으로, 이를 통해 Jupyter Notebook을 빠르게 생성하고 Spark 클러스터에 연결한 후 콘솔에서 Jupyter Notebook 편집기를 열어 쿼리와 코드를 원격으로 실행할 수 있습니다. EMR 노트북은 안정적인 스토리지, 빠른 액세스 및 유연성을 위해 클러스터와 독립적으로 Amazon S3에 저장됩니다. 여러 노트북을 열고, 단일 클러스터에 여러 노트북을 연결하며, 서로 다른 클러스터에서 노트북을 재사용할 수 있습니다.

자세한 내용은 Amazon EMR 관리 안내서에서 EMR 노트북을 참조하세요.

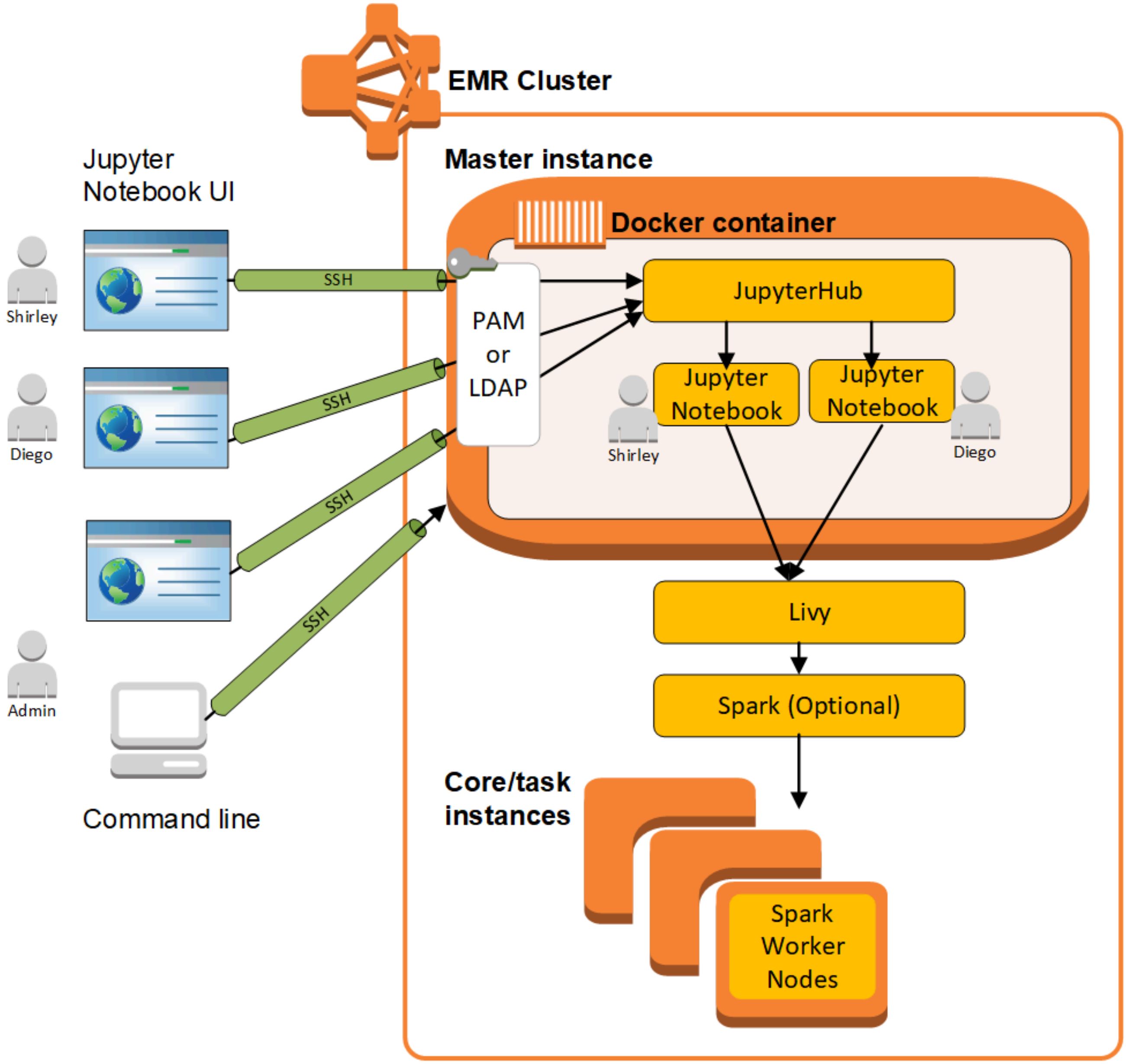
JupyterHub

Jupyter Notebook은 실시간 코드, 수식, 가상화 및 설명 텍스트가 포함된 문서를 생성 및 공유하는 데 사용할 수 있는 오픈 소스 웹 애플리케이션입니다. JupyterHub를 사용하면 단일 사용자 Jupyter Notebook 서버의 여러 인스턴스를 호스팅할 수 있습니다. JupyterHub를 사용하여 클러스터를 생성

하면 Amazon EMR이 클러스터의 프라이머리 노드에 Docker 컨테이너를 생성합니다. JupyterHub, Jupyter에 필요한 모든 구성 요소 및 Sparkmagic이 컨테이너 내에서 실행됩니다.

Sparkmagic은 Jupyter Notebook이 Spark용 REST 서버인 Apache Livy를 통해 Amazon EMR에서 실행되는 Apache Spark와 상호 작용할 수 있도록 하는 커널 라이브러리입니다. Spark 및 Apache Livy는 JupyterHub를 사용하여 클러스터 생성 시 자동으로 설치됩니다. Jupyter용 기본 Python 3 커널은 Sparkmagic에서 제공하는 PySpark 3, PySpark 및 Spark 커널과 함께 사용할 수 있습니다. 이러한 커널을 사용하면 Python 및 Scala를 사용하여 애드혹 Spark 코드 및 대화형 SQL 쿼리를 실행할 수 있습니다. Docker 컨테이너 내에 추가 커널을 수동으로 설치할 수 있습니다. 자세한 내용은 추가 커널 및 라이브러리 설치 단원을 참조하십시오.

다음 다이어그램은 노트북 사용자 및 관리자용 해당 인증 방법을 사용하여 Amazon EMR에 있는 JupyterHub의 구성 요소를 보여줍니다. 자세한 내용은 Jupyter Notebook 사용자 및 관리자 추가 단원을 참조하십시오.



다음 테이블에는 Amazon EMR이 JupyterHub를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 JupyterHub의 버전이 나열되어 있습니다.

이 릴리스에서 JupyterHub와 함께 설치된 구성 요소의 버전은 [릴리스 7.9.0 구성 요소 버전을 참조하세요](#).

emr-7.9.0용 JupyterHub 버전 정보

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-7.9.0	JupyterHub 1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

다음 테이블에는 Amazon EMR이 JupyterHub를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 JupyterHub의 버전이 나열되어 있습니다.

이 릴리스에서 JupyterHub와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 JupyterHub 버전 정보

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.15.0	JupyterHub 1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
		, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

다음 테이블에는 Amazon EMR이 JupyterHub를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 JupyterHub의 버전이 나열되어 있습니다.

이 릴리스에서 JupyterHub와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 JupyterHub 버전 정보

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.36.2	JupyterHub 1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR의 JupyterHub에 포함되어 있는 Python 3 커널은 3.6.4입니다.

`jupyterhub` 컨테이너 내에 설치되어 있는 라이브러리는 Amazon EMR 릴리스 버전 및 Amazon EC2 AMI 버전 간에 다를 수 있습니다.

conda를 사용하여 설치된 라이브러리를 나열하려면

- 마스터 노드 명령줄에서 다음 명령을 실행합니다.

```
sudo docker exec jupyterhub bash -c "conda list"
```

pip를 사용하여 설치된 라이브러리를 나열하려면

- 마스터 노드 명령줄에서 다음 명령을 실행합니다.

```
sudo docker exec jupyterhub bash -c "pip freeze"
```

주제

JupyterHub를 사용하여 클러스터 생성

AWS Management Console AWS Command Line Interface또는 Amazon EMR API를 사용하여 JupyterHub로 Amazon EMR 클러스터를 생성할 수 있습니다. 단계 완료 후 자동으로 종료되는 옵션 (AWS CLI의 `--auto-terminate` 옵션)을 사용하여 클러스터가 생성되지 않도록 해야 합니다. 또한, 관리자 및 노트북 사용자가 클러스터 생성 시 사용된 키 쌍에 액세스할 수 있어야 합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH 보안 인증을 위해 키 페어 사용을 참조하세요.

콘솔을 사용하여 JupyterHub를 포함하는 클러스터 생성

다음 절차를 사용하면 Amazon EMR 콘솔의 고급 옵션을 사용하여 JupyterHub가 설치된 클러스터를 생성합니다.

Amazon EMR 콘솔을 사용하여 JupyterHub가 설치된 Amazon EMR 클러스터를 생성하는 방법

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. 소프트웨어 구성에서 다음을 수행합니다.
   - 릴리스에서는 emr-5.36.2를 선택하고 JupyterHub를 선택합니다.
   - Spark를 사용하는 경우 Glue 데이터 카탈로그를 Spark SQL의 AWS 메타스토어로 사용하려면 Spark 테이블 메타데이터에 사용을 선택합니다. 자세한 내용은 Amazon EMR에서 Spark와 함께 AWS Glue 데이터 카탈로그 사용 단원을 참조하십시오.
   - 소프트웨어 설정 편집에서 구성 입력을 선택하고 값을 지정하거나, S3에서 JSON 로드를 선택하고 JSON 구성 파일을 지정합니다. 자세한 내용은 JupyterHub 구성 단원을 참조하십시오.
4. 단계 추가(선택 사항)에서 클러스터가 생성될 때 실행할 단계를 구성하고, 마지막 단계가 완료된 후 클러스터 자동 종료가 선택되어 있지 않은 상태에서 다음을 선택합니다.
5. 하드웨어 구성 옵션을 선택하고 다음을 선택합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 하드웨어 및 네트워킹 구성을 참조하세요.
6. 일반 클러스터 설정에 대한 옵션을 선택하고 다음을 선택합니다.
7. 키 쌍을 지정한 상태로 보안 옵션을 선택하고 클러스터 생성을 선택합니다.

를 사용하여 JupyterHub로 클러스터 생성 AWS CLI

JupyterHub를 포함하는 클러스터를 시작하려면 `aws emr create-cluster` 명령을 사용하고 `--applications` 옵션에서 `Name=JupyterHub`를 지정합니다. 다음 예에서는 두 개의 EC2 인스턴스(마스터 하나와 코어 인스턴스 하나)를 사용하여 Amazon EMR에서 JupyterHub 클러스터를 실행합니다. 또한 디버깅이 활성화되어 있으며 로그는 `--log-uri`에 의해 지정된 대로 Amazon S3 위치에 저장됩니다. 지정된 키 페어는 클러스터에 있는 Amazon EC2 인스턴스에 액세스할 수 있는 권한만 부여합니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
```

```
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair
```

Amazon EMR에서 JupyterHub를 사용할 때 고려 사항

Amazon EMR에서 JupyterHub를 사용할 때는 다음을 고려합니다.

- **Warning**

  사용자 노트북 및 파일은 마스터 노드에 있는 파일 시스템에 저장됩니다. 이 시스템은 클러스터 종료 시 유지되지 않는 임시 스토리지입니다. 클러스터가 종료되면 백업하지 않은 경우 이 데이터를 잃게 됩니다. `cron` 작업이나 해당 애플리케이션에 적합한 다른 수단을 사용하여 정기적인 백업을 예약해 두는 것이 좋습니다.
  또한, 컨테이너를 다시 시작하는 경우 컨테이너 내 구성 변경 사항도 유지되지 않습니다. 스크립트를 구성하거나 아니면 컨테이너 구성을 자동화하여 보다 쉽게 사용자 지정 구성을 재현할 수 있도록 하는 것이 좋습니다.
- Amazon EMR 보안 구성을 사용하여 설정된 Kerberos 인증은 지원되지 않습니다.
- OAuthenticator는 지원되지 않습니다.

JupyterHub 구성

클러스터 프라이머리 노드에 연결하고 구성 파일을 편집하면 Amazon EMR에 있는 JupyterHub 및 개별 사용자 노트북의 구성을 사용자 지정할 수 있습니다. 값을 변경한 후에는 `jupyterhub` 컨테이너를 다시 시작하십시오.

다음 파일의 속성을 수정하여 JupyterHub 및 개별 Jupyter Notebook을 구성합니다.

- `jupyterhub_config.py` - 기본적으로 이 파일은 프라이머리 노드의 `/etc/jupyter/conf/` 디렉터리에 저장됩니다. 자세한 내용은 JupyterHub 설명서에서 Configuration basics를 참조하세요.
- `jupyter_notebook_config.py` - 이 파일은 기본적으로 `/etc/jupyter/` 디렉터리에 저장되며 `jupyterhub` 컨테이너에 기본값으로 복사됩니다. 자세한 내용은 Jupyter Notebook 설명서의 Config file and command line options를 참조하십시오.

`jupyter-sparkmagic-conf` 구성 분류를 사용하여 Sparkmagic을 사용자 지정할 수도 있습니다. 그러면 Sparkmagic에 맞게 `config.json` 파일의 값이 업데이트됩니다. 사용 가능한 설정에 대한 자세

한 내용은 GitHub의 example_config.json을 참조하십시오. Amazon EMR에서 애플리케이션과 함께 구성 분류를 사용하는 방법에 대한 자세한 내용은 애플리케이션 구성 섹션을 참조하세요.

다음 예제에서는 Sparkmagic 구성 분류 설정에 `MyJupyterConfig.json` 대한 파일을 참조하여 AWS CLI를 사용하여 클러스터를 시작합니다.

Note

가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --use-default-roles --release-label emr-5.14.0 \
--applications Name=Jupyter --instance-type m4.xlarge --instance-count 3 \
--ec2-attributes KeyName=MyKey,SubnetId=subnet-1234a5b6 --configurations file://
MyJupyterConfig.json
```

`MyJupyterConfig.json`의 샘플 콘텐츠는 다음과 같습니다.

```
[
    {
    "Classification":"jupyter-sparkmagic-conf",
    "Properties": {
      "kernel_python_credentials" : "{\"username\":\"diego\",\"base64_password\":
\"mypass\",\"url\":\"http:\/\/localhost:8998\",\"auth\":\"None\"}"
      }
    }
]
```

Note

Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

Amazon S3에서 노트북의 지속성 구성

Amazon EMR에서 JupyterHub 클러스터를 구성하여 사용자가 저장한 노트북이 클러스터 EC2 인스턴스의 임시 스토리지 외부에 있는 Amazon S3에 지속되도록 할 수 있습니다.

클러스터를 생성할 때 `jupyter-s3-conf` 구성 분류를 사용하여 Amazon S3 지속성을 지정합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

`s3.persistence.enabled` 속성을 사용하여 Amazon S3 지속성을 활성화하는 것 외에도 Amazon S3에서 `s3.persistence.bucket` 속성을 사용하여 노트북을 저장하는 버킷을 지정합니다. 각 사용자의 노트북은 지정된 버킷의 `jupyter/`*`jupyterhub-user-name`* 폴더에 저장됩니다. 버킷은 Amazon S3에 이미 있어야 하며 클러스터를 생성할 때 지정하는 EC2 인스턴스 프로필의 역할에는 버킷에 대한 사용 권한이 있어야 합니다(기본 역할: `EMR_EC2_DefaultRole`). 자세한 내용은 [AWS 서비스에 대한 Amazon EMR 권한에 대한 IAM 역할 구성을 참조하세요](#).

동일한 구성 분류 속성을 사용하여 새 클러스터를 시작하면 사용자는 저장된 위치의 콘텐츠로 노트북을 열 수 있습니다.

Amazon S3를 활성화한 상태에서 노트북의 모듈로 파일을 가져오면 파일이 Amazon S3에 업로드됩니다. Amazon S3 지속성을 활성화하지 않고 파일을 가져오면 파일이 JupyterHub 컨테이너에 업로드됩니다.

다음 예제에서는 Amazon S3 지속성을 활성화합니다. 노트북은 각 사용자의 `s3://MyJupyterBackups/jupyter/`*`jupyterhub-user-name`* 폴더에 저장됩니다. 여기서 *`jupyterhub-user-name`*은 사용자 이름(예: `diego`)입니다.

```
[
    {
        "Classification": "jupyter-s3-conf",
        "Properties": {
            "s3.persistence.enabled": "true",
            "s3.persistence.bucket": "MyJupyterBackups"
        }
    }
]
```

프라이머리 노드 및 노트북 서버에 연결

JupyterHub 관리자 및 노트북 사용자는 SSH 터널을 사용한 다음 마스터 노드의 JupyterHub에 의해 지원되는 웹 인터페이스에 연결하여 클러스터 마스터 노드에 연결해야 합니다. SSH 터널을 구성하고 터

널을 사용하여 웹 연결을 프록시하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 [클러스터에 연결](#)을 참조하세요.

기본적으로 Amazon EMR의 JupyterHub는 프라이머리 노드의 포트 9443을 통해 사용할 수 있습니다. 내부 JupyterHub 프록시는 포트 9443을 통해 노트북 인스턴스도 지원합니다. JupyterHub 및 Jupyter 웹 인터페이스는 다음 패턴으로 URL을 사용하여 액세스할 수 있습니다.

https://*MasterNodeDNS*:9443

`c.JupyterHub.port` 파일에서 `jupyterhub_config.py` 속성을 사용하여 다른 포트를 지정할 수 있습니다. 자세한 내용은 JupyterHub 설명서에서 [Networking basics](#)를 참조하세요.

기본적으로 Amazon EMR에 설치된 JupyterHub에서는 HTTPS를 사용한 SSL 암호화를 위해 자체 서명된 인증서를 사용합니다. 사용자는 연결 시 자체 서명된 인증서를 신뢰하라는 메시지를 받게 됩니다. 신뢰된 인증서 및 키를 자체적으로 사용할 수 있습니다. 마스터 노드의 `server.crt`에 있는 기본 인증서 파일, `server.key` 및 키 파일 `/etc/jupyter/conf/`를 자체 인증서 및 키 파일로 바꾸십시오. `c.JupyterHub.ssl_key` 파일에 있는 `c.JupyterHub.ssl_cert` 및 `jupyterhub_config.py` 속성을 사용하여 SSL 재료를 지정하십시오. 자세한 내용은 JupyterHub 설명서에서 [Security settings](#)를 참조하세요. `jupyterhub_config.py`를 업데이트한 후 컨테이너를 다시 시작하십시오.

JupyterHub 구성 및 관리

JupyterHub 및 관련 구성 요소는 Ubuntu 운영 체제를 실행하는 `jupyterhub`라는 Docker 컨테이너 내에서 실행됩니다. 컨테이너 내에서 실행 중인 구성 요소를 관리하는 방법에는 여러 가지가 있습니다.

> **Warning**
>
> 컨테이너 내에서 수행하는 사용자 지정 작업은 컨테이너를 다시 시작하는 경우 유지되지 않을 수 있습니다. 스크립트를 구성하거나 아니면 컨테이너 구성을 자동화하여 보다 쉽게 사용자 지정 구성을 재현할 수 있도록 하는 것이 좋습니다.

명령줄을 사용하여 관리

SSH를 사용하여 마스터 노드에 연결되면 Docker 명령줄 인터페이스(CLI)를 사용하고 이름(`jupyterhub`)이나 ID로 컨테이너를 지정하여 명령을 실행할 수 있습니다. 예를 들어, `sudo docker exec jupyterhub *command*`는 컨테이너 내에서 실행 중인 애플리케이션이나 운영 체제에 의해 인식되는 명령을 실행합니다. 이 명령을 사용하면 운영 체제에 사용자를 추가하고 추가 애플리케이션 및 라이브러리를 Docker 컨테이너 내에 설치할 수 있습니다. 예를 들어, 기본 컨테이너 이미지에는 패키

지 설치 시 사용할 Conda가 들어 있으므로 마스터 노드 명령줄에서 다음 명령을 실행하면 애플리케이션, Keras를 컨테이너 내에 설치할 수 있습니다.

```
sudo docker exec jupyterhub conda install keras
```

단계를 제출하여 관리

단계는 작업을 클러스터에 제출하는 방법입니다. 단계는 클러스터를 실행할 때 제출할 수도 있고, 실행 중인 클러스터에 제출할 수도 있습니다. 명령줄에서 실행하는 명령은 `command-runner.jar`를 사용하여 단계로 제출될 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 CLI 및 콘솔을 사용하여 단계 작업 및 Amazon EMR 클러스터에서 명령 및 스크립트 실행 섹션을 참조하세요.

예를 들어 로컬 컴퓨터에서 다음 AWS CLI 명령을 사용하여 이전 예제에서 마스터 노드의 명령줄에서 수행한 것과 동일한 방식으로 Keras를 설치할 수 있습니다.

```
aws emr add-steps --cluster-id MyClusterID --steps Name="Command
 Runner",Jar="command-runner.jar",Args="/usr/bin/sudo","/usr/bin/
docker","exec","jupyterhub","conda","install","keras"
```

또한 일련의 단계를 스크립트로 구성하고 스크립트를 Amazon S3에 업로드한 다음 `script-runner.jar`을 사용하여 클러스터를 생성하거나 스크립트를 단계로 추가할 때 스크립트를 실행할 수 있습니다. 자세한 내용은 Amazon EMR 클러스터에서 명령 및 스크립트 실행 단원을 참조하십시오. 예제는 the section called “예제: 여러 사용자를 추가하는 Bash 스크립트” 섹션을 참조하세요.

REST API를 사용하여 관리

Jupyter, JupyterHub 및 JupyterHub용 HTTP 프록시는 요청을 보내는 데 사용할 수 있는 REST API를 제공합니다. JupyterHub에 요청을 보내려면, 요청이 있는 API 토큰을 전달해야 합니다. 마스터 노드 명령줄에서 `curl` 명령을 사용하면 REST 명령을 실행할 수 있습니다. 자세한 정보는 다음 자료를 참조하세요.

- API 토큰 생성에 대한 지침에 들어 있는 JupyterHub용 설명서의 Using JupyterHub's REST API(JupyterHub의 REST API 사용)
- GitHub의 Jupyter Notebook Server API
- GitHub의 configurable-http-proxy

다음은 JupyterHub용 REST API를 사용하여 사용자 목록을 가져오는 방법을 보여주는 예제입니다. 이 명령은 이전에 생성된 관리 토큰을 전달하고 기본 포트 9443를 JupyterHub에 사용하여 보다 쉽게 볼 수 있도록 결과를 jq에 보냅니다.

```
curl -XGET -s -k https://$HOST:9443/hub/api/users \
-H "Authorization: token $admin_token" | jq .
```

Jupyter Notebook 사용자 및 관리자 추가

두 가지 방법 중 하나를 사용하면 사용자가 JupyterHub에 인증하여 노트북을 생성할 수 있으며, 선택적으로 JupyterHub를 관리할 수도 있습니다. 가장 쉬운 방법은 JupyterHub의 플러그형 인증 모듈(PAM)을 사용하는 것입니다. 또한 Amazon EMR에 설치된 JupyterHub는 JupyterHub용 LDAP 인증자 플러그인을 지원하여 Microsoft Active Directory 서버와 같은 LDAP 서버로부터 사용자 ID를 가져옵니다. 각 인증 방법을 사용하여 사용자를 추가하는 방법에 대한 지침과 예제는 본 단원에서 제공됩니다.

Amazon EMR의 JupyterHub에는 관리자 권한이 있는 기본 사용자가 있습니다. 사용자 이름은 `jovyan`이고, 암호는 `jupyter`입니다. 이 사용자를 관리자 권한이 있는 다른 사용자로 바꾸는 것이 매우 권장됩니다. 클러스터를 생성할 때 단계를 사용하거나 클러스터를 실행할 때 마스터 노드에 연결하여 그렇게 할 수 있습니다.

주제

- PAM 인증 사용
- LDAP 인증 사용
- 사용자 위장

PAM 인증 사용

Amazon EMR에 설치된 JupyterHub에 PAM 사용자를 생성하는 방법은 두 단계의 프로세스로 이루어집니다. 첫 번째 단계는 마스터 노드의 `jupyterhub` 컨테이너에서 실행 중인 운영 체제에 사용자를 추가하는 것과 해당 사용자를 각 사용자의 홈 디렉터리에 추가하는 것입니다. 두 번째 단계는 이러한 운영 체제 사용자를 JupyterHub 사용자로 추가하는 것입니다. JupyterHub에서는 이 프로세스를 화이트리스팅이라고 합니다. 추가된 JupyterHub 사용자는 JupyterHub URL에 연결한 다음, 액세스를 위하 자신의 운영 체제 자격 증명을 제공하면 됩니다.

사용자가 로그인하면 JupyterHub는 해당 사용자용 노트북 서버 인스턴스를 엽니다. 이 인스턴스는 마스터 노드에 있는 사용자의 홈 디렉터리에 저장됩니다(`/var/lib/jupyter/home/`*`username`*). 노트북 서버 인스턴스가 없는 경우, JupyterHub에서 사용자의 홈 디렉터리에 노트북 인스턴스를 만듭니다.

다음 단원에서는 여러 사용자를 추가하는 기초적인 bash 스크립트를 따라 운영 체제 및 JupyterHub에 사용자를 개별적으로 추가하는 방법을 보여줍니다.

컨테이너에 운영 체제 사용자 추가

다음 예제에서는 먼저 컨테이너 내에서 useradd 명령을 사용하여 단일 사용자인 diego를 추가한 다음 해당 사용자용 홈 디렉터리를 만듭니다. 두 번째 명령에서는 chpasswd를 사용하여 이 사용자를 위한 diego 암호를 설정합니다. 명령은 SSH를 사용하여 연결되어 있는 동안 마스터 노드 명령줄에서 실행됩니다. 단계를 제출하여 관리의 앞부분에 설명되어 있는 대로 단계를 사용하여 이 명령을 실행할 수도 있습니다.

```
sudo docker exec jupyterhub useradd -m -s /bin/bash -N diego
sudo docker exec jupyterhub bash -c "echo diego:diego | chpasswd"
```

JupyterHub 사용자 추가

JupyterHub의 Admin 패널이나 REST API를 사용하면 사용자 및 관리자를 추가하거나 사용자만 추가할 수 있습니다.

JupyterHub의 Admin 패널을 사용하면 사용자 및 관리자를 추가하는 방법

1. SSH를 사용하여 마스터 노드에 연결하고 관리자 권한이 있는 ID를 사용하여 https://*MasterNodeDNS*:9443에 로그인합니다.
2. 제어판, Admin을 선택합니다.
3. 사용자, 사용자 추가를 선택하거나 Admin, Add Admins(Admin 추가)를 선택합니다.

REST API를 사용하여 사용자를 추가하려면

1. SSH를 사용하여 마스터 노드에 연결하고 마스터 노드에서 다음 명령을 사용하거나 하나의 단계로 명령을 실행합니다.
2. API 요청을 만들 관리자 토큰을 획득하고 다음 단계의 *AdminToken*을 해당 토큰으로 바꿉니다.
3. 다음 명령을 사용하되, *UserName*을 컨테이너 내에서 생성된 운영 체제 사용자로 바꿉니다.

```
curl -XPOST -H "Authorization: token AdminToken" "https://$(hostname):9443/hub/api/
users/UserName
```

> **Note**
>
> JupyterHub 웹 인터페이스에 처음 로그인하면 자동으로 관리자가 아닌 JupyterHub 사용자로 추가됩니다.

예제: 여러 사용자를 추가하는 Bash 스크립트

다음 샘플 bash 스크립트는 이 단원의 앞의 단계와 함께 연결되어 여러 JupyterHub 사용자를 생성합니다. 이 스크립트는 프라이머리 노드에서 직접 실행할 수도 있고, Amazon S3에 업로드한 다음 하나의 단계로 실행할 수도 있습니다.

이 스크립트는 먼저 사용자 이름 배열을 설정한 다음, `jupyterhub token` 명령을 사용하여 기본 관리자인 jovyan용 API 토큰을 생성합니다. 그런 다음 해당 사용자용 `jupyterhub` 컨테이너에 운영 체제 사용자를 생성합니다. 이때 최초 암호는 각 사용자 이름과 동일하게 지정됩니다. 마지막으로, REST API 작업을 호출하여 JupyterHub에 각 사용자를 생성합니다. 또한 스크립트의 앞부분에서 생성한 토큰을 전달하고 REST 응답을 `jq`에 보내 보다 쉽게 볼 수 있도록 해줍니다.

```
# Bulk add users to container and JupyterHub with temp password of username
set -x
USERS=(shirley diego ana richard li john mary anaya)
TOKEN=$(sudo docker exec jupyterhub /opt/conda/bin/jupyterhub token jovyan | tail -1)
for i in "${USERS[@]}";
do
   sudo docker exec jupyterhub useradd -m -s /bin/bash -N $i
   sudo docker exec jupyterhub bash -c "echo $i:$i | chpasswd"
   curl -XPOST --silent -k https://$(hostname):9443/hub/api/users/$i \
 -H "Authorization: token $TOKEN" | jq
done
```

이 스크립트를 Amazon S3의 위치(예: `s3://amzn-s3-demo-bucket/createjupyterusers.sh`)에 저장합니다. 그런 다음 `script-runner.jar`을 사용하여 이를 하나의 단계로 실행하면 됩니다.

예제: 클러스터 생성 시 스크립트 실행(AWS CLI)

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/MyJupyterClusterLogs
 \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair \
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-
s3-demo-bucket/createjupyterusers.sh"]
```

기존 클러스터에서 스크립트 실행(AWS CLI)

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr add-steps --cluster-id j-XXXXXXXX --steps Type=CUSTOM_JAR,\
Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-
s3-demo-bucket/createjupyterusers.sh"]
```

LDAP 인증 사용

LDAP(Lightweight Directory Access Protocol)는 Active Directory 또는 OpenLDAP 서버와 같은 LDAP 호환 디렉터리 서비스 공급자에 저장된 사용자 및 컴퓨터와 같은 리소스에 해당하는 객체를 쿼리 및 수정하기 위한 애플리케이션 프로토콜입니다. Amazon EMR에서 JupyterHub와 함께 JupyterHub용 LDAP 인증자 플러그인을 사용하여 사용자 인증에 LDAP를 사용할 수 있습니다. 이 플러그인은 LDAP 사용자용 로그인 세션을 처리하고 사용자 정보를 Jupyter에 제공합니다. 이렇게 하면 사용자가 LDAP 호환 서버에 저장된 자신의 ID에 대한 자격 증명을 사용하여 JupyterHub 및 노트북에 연결할 수 있습니다.

이 단원의 단계에서는 다음 단계를 통해 JupyterHub용 LDAP 인증자 플러그인을 사용하여 LDAP를 설정 및 활성화하는 방법을 안내합니다. 마스터 노드 명령줄에 연결되어 있는 동안 단계를 수행해야 합니다. 자세한 내용은 프라이머리 노드 및 노트북 서버에 연결 단원을 참조하십시오.

1. 호스트 IP 주소, 포트, 바인딩 이름 등과 같은 LDAP 서버에 대한 정보를 사용하여 LDAP 구성 파일을 생성합니다.

2. JupyterHub용 LDAP 인증자 플러그인을 활성화하도록 `/etc/jupyter/conf/jupyterhub_config.py`를 수정합니다.
3. `jupyterhub` 컨테이너 내에서 LDAP를 구성하는 스크립트를 생성하고 실행합니다.
4. 사용자용 LDAP를 조회한 다음, 해당 사용자를 위해 컨테이너 내에 홈 디렉터리를 생성합니다. JupyterHub를 사용하려면 노트북을 호스팅하는 홈 디렉터리가 있어야 합니다.
5. JupyterHub를 다시 시작하는 스크립트 실행

> **Important**
>
> LDAP를 설정하기 전에, 네트워크 인프라를 테스트하여 LDAP 서버 및 클러스터 마스터 노드가 필요 시 통신할 수 있도록 합니다. TLS는 일반적으로 일반 TCP 연결을 통해 포트 389를 사용합니다. LDAP 연결에서 SSL을 사용하는 경우 SSL용 잘 알려진 TCP 포트는 636입니다.

LDAP 구성 파일 생성

아래 예에서는 다음 자리 표시자 구성 값을 사용합니다. 이를 자신의 구현 상황에 맞는 파라미터로 바꿉니다.

- LDAP 서버는 버전 3을 실행하고 있으며 포트 389에서 사용할 수 있습니다. 이 포트는 LDAP용 표준 비SSL 포트입니다.
- 기본 고유 이름(DN)은 `dc=example, dc=org`입니다.

텍스트 편집기를 사용하여 다음과 유사한 내용으로 [ldap.conf](#) 파일을 생성합니다. LDAP 구현 상황에 적절한 값을 사용합니다. *`host`*를 LDAP 서버의 IP 주소 또는 확인 가능한 호스트 이름으로 바꿉니다.

```
base dc=example,dc=org
uri ldap://host
ldap_version 3
binddn cn=admin,dc=example,dc=org
bindpw admin
```

JupyterHub용 LDAP 인증자 플러그인 활성화

텍스트 편집기를 사용하여 `/etc/jupyter/conf/jupyterhub_config.py` 파일을 수정하고 다음과 유사하게 [ldapauthenticator](#) 속성을 추가합니다. *`host`*를 LDAP 서버의 IP 주소 또는 확인 가능

한 호스트 이름으로 바꿉니다. 이 예에서는 사용자 객체가 *people*이라는 조직 단위(OU) 내에 있으며 ldap.conf를 사용하여 앞에서 설정한 고유 이름 구성 요소를 사용한다고 가정합니다.

```
c.JupyterHub.authenticator_class = 'ldapauthenticator.LDAPAuthenticator'
c.LDAPAuthenticator.use_ssl = False
c.LDAPAuthenticator.server_address = 'host'
c.LDAPAuthenticator.bind_dn_template = 'cn={username},ou=people,dc=example,dc=org'
```

컨테이너 내에서 LDAP 구성

텍스트 편집기를 사용하여 다음 콘텐츠가 포함된 bash 스크립트를 생성합니다.

```
#!/bin/bash

# Uncomment the following lines to install LDAP client libraries only if
# using Amazon EMR release version 5.14.0. Later versions install libraries by default.
# sudo docker exec jupyterhub bash -c "sudo apt-get update"
# sudo docker exec jupyterhub bash -c "sudo apt-get -y install libnss-ldap libpam-ldap
 ldap-utils nscd"

# Copy ldap.conf
sudo docker cp ldap.conf jupyterhub:/etc/ldap/
sudo docker exec jupyterhub bash -c "cat /etc/ldap/ldap.conf"

# configure nss switch
sudo docker exec jupyterhub bash -c "sed -i 's/\(^passwd.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "sed -i 's/\(^group.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "sed -i 's/\(^shadow.*\)/\1 ldap/g' /etc/
nsswitch.conf"
sudo docker exec jupyterhub bash -c "cat /etc/nsswitch.conf"

# configure PAM to create home directories
sudo docker exec jupyterhub bash -c "echo 'session required        pam_mkhomedir.so
 skel=/etc/skel umask=077' >> /etc/pam.d/common-session"
sudo docker exec jupyterhub bash -c "cat /etc/pam.d/common-session"

# restart nscd service
sudo docker exec jupyterhub bash -c "sudo service nscd restart"

# Test
sudo docker exec jupyterhub bash -c "getent passwd"
```

```
# Install ldap plugin
sudo docker exec jupyterhub bash -c "pip install jupyterhub-ldapauthenticator"
```

스크립트를 마스터 노드에 저장한 다음 이를 마스터 노드 명령줄에서 실행합니다. 예를 들어, `configure_ldap_client.sh`라고 저장한 스크립트를 사용하여 파일을 실행 파일로 만듭니다.

```
chmod +x configure_ldap_client.sh
```

그런 다음 스크립트를 실행합니다.

```
./configure_ldap_client.sh
```

Active Directory에 속성 추가

데이터베이스에서 각 사용자를 찾고 적절한 항목을 만들려면, JupyterHub 도커 컨테이너에 Active Directory의 해당 사용자 객체에 대한 다음 UNIX 속성이 필요합니다. 자세한 내용은 [Clarification regarding the status of identity management for Unix (IDMU) and NIS server role in Windows Server 2016 technical preview and beyond]() 기사의 How do I continue to edit the GID/UID RFC 2307 attributes now that the Unix Attributes Plug-in is no longer available for the Active Directory Users and Computers MMC snap-in? 섹션을 참조하세요.

- `homeDirectory`

  사용자 홈 디렉터리의 위치이며 보통 `/home/username`입니다.

- `gidNumber`

  다른 사용자가 아직 사용하지 않은 60000보다 큰 값입니다. 사용 중인 GID는 `etc/passwd` 파일에서 확인하십시오.

- `uidNumber`

  다른 그룹이 아직 사용하지 않은 60000보다 큰 값입니다. 사용 중인 UID는 `etc/group` 파일에서 확인하십시오.

- `uid`

  *username*과 같습니다.

사용자 홈 디렉터리 생성

JupyterHub에서는 LDAP 사용자를 인증하고 인스턴스 데이터를 저장하기 위해 컨테이너 내에 홈 디렉터리가 있어야 합니다. 다음 예는 LDAP 디렉터리에 있는 두 명의 사용자 shirley와 diego를 보여줍니다.

첫 번째 단계는 다음 예에 표시된 대로 ldapsearch를 사용하여 각 사용자의 사용자 ID 및 그룹 ID 정보를 LDAP 서버에 조회하는 것입니다. 이때 *host*는 LDAP 서버의 IP 주소 또는 확인 가능한 호스트 이름으로 바꿉니다.

```
ldapsearch -x -H ldap://host \
 -D "cn=admin,dc=example,dc=org" \
 -w admin \
 -b "ou=people,dc=example,dc=org" \
 -s sub \
 "(objectclass=*)" uidNumber gidNumber
```

`ldapsearch`명령은 shirley와 diego 사용자를 위해 다음과 유사하게 표시되는 LDIF 형식 응답을 반환합니다.

```
# extended LDIF

# LDAPv3
# base <ou=people,dc=example,dc=org> with scope subtree
# filter: (objectclass=*)
# requesting: uidNumber gidNumber sn

# people, example.org
dn: ou=people,dc=example,dc=org

# diego, people, example.org
dn: cn=diego,ou=people,dc=example,dc=org
sn: B
uidNumber: 1001
gidNumber: 100

# shirley, people, example.org
dn: cn=shirley,ou=people,dc=example,dc=org
sn: A
uidNumber: 1002
gidNumber: 100
```

```
# search result
search: 2
result: 0 Success

# numResponses: 4
# numEntries: 3
```

응답에서 획득한 정보로 컨테이너 내 명령을 실행하여 각 사용자의 일반 이름(cn)용 홈 디렉터리를 생성합니다. `uidNumber` 및 `gidNumber`를 사용하여 해당 사용자에 대한 홈 디렉터리용 소유권을 수정합니다. 다음 예시 명령은 사용자 *shirley*에 대해 이 작업을 수행합니다.

```
sudo docker container exec jupyterhub bash -c "mkdir /home/shirley"
sudo docker container exec jupyterhub bash -c "chown -R $uidNumber /home/shirley"
sudo docker container exec jupyterhub bash -c "sudo chgrp -R $gidNumber /home/shirley"
```

> **Note**
>
> JupyterHub의 LDAP 인증자는 로컬 사용자 생성을 지원하지 않습니다. 자세한 내용은 [LDAP authenticator configuration note on local user creation]을 참조하세요.
> 로컬 사용자를 수동으로 생성하려면 다음 명령을 사용합니다.
>
> ```
> sudo docker exec jupyterhub bash -c "echo 'shirley:x:$uidNumber:$gidNumber::/
> home/shirley:/bin/bash' >> /etc/passwd"
> ```

Jupyterhub 컨테이너 다시 시작

`jupyterhub` 컨테이너를 다시 시작하려면 다음 명령을 실행합니다.

```
sudo docker stop jupyterhub
sudo docker start jupyterhub
```

사용자 위장

Jupyter Notebook 내에서 실행되는 Spark 작업은 Amazon EMR에서 실행되는 동안 여러 애플리케이션을 통과합니다. 예를 들어 사용자가 Jupyter 내에서 실행하는 PySpark3 코드를 Sparkmagic이 수신하며, 이 Sparkmagic은 수신한 PySpark3 코드를 HTTP POST 요청을 사용하여 Livy로 제출합니다. 이때 Livy는 YARN을 사용해 클러스터에서 실행할 Spark 작업을 생성합니다.

기본적으로 이런 방식으로 제출된 YARN 작업은 최초로 작업을 시작한 사용자와 상관없이 사용자 `livy`로서 실행됩니다. 사용자 가장을 설정함으로써 노트북 사용자의 사용자 ID 또한 YARN 작업과 관련된 사용자가 되도록 설정할 수 있습니다. 작업을 사용자 `shirley`와 연결된 `diego` 및 `livy` 모두로 시작한다기보다는, 각 사용자가 시작하는 작업이 각각 `shirley` 및 `diego`와 연결되는 것입니다. 이렇게 하면 Jupyter 사용량을 감사하고 조직 내에서 애플리케이션을 관리할 수 있습니다.

이러한 구성은 Sparkmagic의 Livy에 대한 호출이 인증되지 않는 경우에만 지원됩니다. 하둡 애플리케이션과 Livy 사이에 인증 또는 프록시 설정 계층을 제공하는 애플리케이션(Apache Knox Gateway 등)은 지원되지 않습니다. 이 섹션의 사용자 가장을 구성하는 단계는 JupyterHub와 Livy가 동일한 마스터 노드에서 실행되는 것을 가정했습니다. 애플리케이션에 별도의 클러스터가 있는 경우 Livy 마스터 노드에 HDFS 디렉터리가 생성되도록 3단계: 사용자를 위한 HDFS 홈 디렉터리 생성을 수정할 필요가 있습니다.

사용자 가장 구성 단계

- 1단계: Livy 구성
- 2단계: 사용자 추가
- 3단계: 사용자를 위한 HDFS 홈 디렉터리 생성

1단계: Livy 구성

다음 예와 같이 클러스터를 생성하여 Livy 사용자 가장을 활성화할 때 `livy-conf` 및 `core-site` 구성 분류를 사용합니다. 구성 분류를 JSON으로 저장한 다음 클러스터를 생성할 때 참조하거나 구성 분류 인라인을 지정합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

```
[
  {
    "Classification": "livy-conf",
    "Properties": {
      "livy.impersonation.enabled": "true"
    }
  },
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.proxyuser.livy.groups": "*",
      "hadoop.proxyuser.livy.hosts": "*"
    }
  }
]
```

2단계: 사용자 추가

PAM 또는 LDAP를 사용하여 JupyterHub 사용자를 추가합니다. 자세한 내용은 PAM 인증 사용 및 LDAP 인증 사용 섹션을 참조하세요.

3단계: 사용자를 위한 HDFS 홈 디렉터리 생성

마스터 노드에 연결하여 사용자를 생성했습니다. 아직 마스터 노드에 연결된 상태에서 아래 콘텐츠를 복사하여 스크립트 파일에 저장합니다. 스크립트는 마스터 노드의 각 JupyterHub 사용자에게 HDFS 홈 디렉터리를 생성합니다. 스크립트는 사용자가 기본 관리자 사용자 ID, *jovyan*을 사용하는 것으로 가정합니다.

```
#!/bin/bash

CURL="curl --silent -k"
HOST=$(curl -s http://169.254.169.254/latest/meta-data/local-hostname)

admin_token() {
    local user=jovyan
    local pwd=jupyter
    local token=$($CURL https://$HOST:9443/hub/api/authorizations/token \
        -d "{\"username\":\"$user\", \"password\":\"$pwd\"}" | jq ".token")
    if [[ $token != null ]]; then
        token=$(echo $token | sed 's/"//g')
    else
        echo "Unable to get Jupyter API Token."
        exit 1
    fi
    echo $token
}

# Get Jupyter Admin token
token=$(admin_token)

# Get list of Jupyter users
users=$(curl -XGET -s -k https://$HOST:9443/hub/api/users \
 -H "Authorization: token $token" | jq '.[].name' | sed 's/"//g')

# Create HDFS home dir
for user in ${users[@]};
do
 echo "Create hdfs home dir for $user"
 hadoop fs -mkdir /user/$user
```

```
 hadoop fs -chmod 777 /user/$user
done
```

추가 커널 및 라이브러리 설치

Amazon EMR에서 JupyterHub를 포함하는 클러스터를 생성하는 경우, Jupyter용 기본 Python 3 커널 및 Sparkmagic용 PySpark 및 Spark 커널이 Docker 컨테이너에 설치됩니다. 추가 커널을 설치할 수 있으며, 추가 라이브러리 및 패키지를 설치한 다음 해당 셸을 위해 이를 가져올 수도 있습니다.

커널 설치

커널은 Docker 컨테이너 내에 설치됩니다. 이를 수행하는 가장 쉬운 방법은 설치 명령이 포함된 bash 스크립트를 작성하고, 이를 마스터 노드에 저장한 다음, `sudo docker exec jupyterhub` *`script_name`* 명령을 사용하여 `jupyterhub` 컨테이너 내에서 스크립트를 실행하는 것입니다. 다음 스크립트 예에서는 커널을 설치한 다음 해당 커널용 라이브러리 몇 개를 마스터 노드에 설치하여 사용자가 나중에 Jupyter에서 커널을 사용하여 라이브러리를 가져올 수 있도록 해줍니다.

```
#!/bin/bash

# Install Python 2 kernel
conda create -n py27 python=2.7 anaconda
source /opt/conda/envs/py27/bin/activate
apt-get update
apt-get install -y gcc
/opt/conda/envs/py27/bin/python -m pip install --upgrade ipykernel
/opt/conda/envs/py27/bin/python -m ipykernel install

# Install libraries for Python 2
/opt/conda/envs/py27/bin/pip install paramiko nltk scipy numpy scikit-learn pandas
```

컨테이너 내에 커널 및 라이브러리를 설치하려면 마스터 노드에 대한 터미널 연결을 열고, 스크립트를 `/etc/jupyter/install_kernels.sh`에 저장한 다음, 마스터 노드 명령줄에서 다음 명령을 실행합니다.

```
sudo docker exec jupyterhub bash /etc/jupyter/install_kernels.sh
```

라이브러리 사용 및 추가 라이브러리 설치

핵심 기계 학습 세트 및 Python 3용 데이터 과학 라이브러리는 Amazon EMR에 JupyterHub와 함께 사전 설치되어 있습니다. `sudo docker exec jupyterhub bash -c "conda list"` 및 `sudo docker exec jupyterhub bash -c "pip freeze"`를 사용할 수 있습니다.

Spark 작업에서 작업자 노드에 있는 라이브러리를 필요로 하는 경우, 부트스트랩 작업을 사용하여 클러스터 생성 시 라이브러리를 설치하는 스크립트를 실행하는 것이 좋습니다. 부트스트랩 작업은 클러스터 생성 프로세서 중에 모든 클러스터 노드에서 실행되므로 설치 작업을 간소화해 줍니다. 클러스터 실행 후 코어/작업자 노드에 라이브러리를 설치하는 경우, 작업이 더 복잡합니다. 이 단원에서는 이러한 라이브러리 설치 방법을 보여주는 Python 예제 프로그램이 제공됩니다.

이 섹션에 표시된 부트스트랩 작업 및 Python 프로그램에서는 Amazon S3에 저장된 bash 스크립트를 사용하여 모든 노드에 라이브러리를 설치합니다.

다음 예제에서 참조되는 스크립트는 `pip`를 사용하여 Python 3 커널용 paramiko, nltk, scipy, scikit-learn, pandas를 설치합니다.

```
#!/bin/bash

sudo python3 -m pip install boto3 paramiko nltk scipy scikit-learn pandas
```

스크립트를 생성했으면 이를 Amazon S3의 위치(예: `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`)에 업로드합니다. 부트스트랩 작업이나 Python 프로그램에서 사용할 수 있도록 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 [객체 업로드]를 참조하세요.

를 사용하여 클러스터를 생성할 때 모든 노드에 라이브러리를 설치하는 부트스트랩 작업을 지정하려면 AWS CLI

1. 앞의 예제와 유사한 스크립트를 생성하고 이를 Amazon S3의 위치에 저장합니다. 이 예에서는 `s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh`를 사용합니다.
2. JupyterHub를 포함하는 클러스터를 생성하고 `--bootstrap-actions` 옵션의 `Path` 인수를 사용하여 다음 예제에 표시된 대로 스크립트 위치를 지정합니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name="MyJupyterHubCluster" --release-label emr-5.36.2 \
--applications Name=JupyterHub --log-uri s3://amzn-s3-demo-bucket/
MyJupyterClusterLogs \
--use-default-roles --instance-type m5.xlarge --instance-count 2 --ec2-attributes
 KeyName=MyKeyPair \
--bootstrap-actions Path=s3://amzn-s3-demo-bucket/install-my-jupyter-
libraries.sh,Name=InstallJupyterLibs
```

콘솔을 사용하여 클러스터 생성 시 모든 노드에 라이브러리를 설치하는 부트스트랩 작업을 지정하려면

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. 소프트웨어 및 단계 및 하드웨어의 설정을 애플리케이션에 적절하게 지정합니다.
4. 일반 클러스터 설정 화면에서 부트스트랩 작업을 확장합니다.
5. 부트스트랩 작업 추가에서 사용자 지정 작업, 구성 및 추가를 선택합니다.
6. 명칭에 기억하기 쉬운 명칭을 입력합니다. 스크립트 위치에, 스크립트가 있는 Amazon S3의 위치를 입력합니다(이 예제에서는 s3://amzn-s3-demo-bucket/install-my-jupyter-libraries.sh를 사용함). 선택적 인수를 비워 두고 추가를 선택합니다.
7. 클러스터에 대한 다른 설정을 지정하고 다음을 선택합니다.
8. 보안 설정을 지정하고 클러스터 생성을 선택합니다.

Example 실행 중인 클러스터의 코어 노드에 라이브러리 설치

Jupyter 내의 마스터 노드에 라이브러리를 설치했으면 다양한 방법으로 실행 중인 코어 노드에 라이브러리를 설치할 수 있습니다. 다음 예제에서는 로컬 컴퓨터에서 실행하도록 작성된 Python 프로그램을 보여줍니다. Python 프로그램을 로컬에서 실행하면 `AWS-RunShellScript`의를 사용하여 클러스터의 코어 노드에 라이브러리를 설치하는이 섹션의 앞부분에 표시된 예제 스크립트를 AWS Systems Manager 실행합니다.

```
import argparse
import time
import boto3
```

```
def install_libraries_on_core_nodes(cluster_id, script_path, emr_client, ssm_client):
    """
    Copies and runs a shell script on the core nodes in the cluster.

    :param cluster_id: The ID of the cluster.
    :param script_path: The path to the script, typically an Amazon S3 object URL.
    :param emr_client: The Boto3 Amazon EMR client.
    :param ssm_client: The Boto3 AWS Systems Manager client.
    """
    core_nodes = emr_client.list_instances(
        ClusterId=cluster_id, InstanceGroupTypes=["CORE"]
    )["Instances"]
    core_instance_ids = [node["Ec2InstanceId"] for node in core_nodes]
    print(f"Found core instances: {core_instance_ids}.")

    commands = [
        # Copy the shell script from Amazon S3 to each node instance.
        f"aws s3 cp {script_path} /home/hadoop",
        # Run the shell script to install libraries on each node instance.
        "bash /home/hadoop/install_libraries.sh",
    ]
    for command in commands:
        print(f"Sending '{command}' to core instances...")
        command_id = ssm_client.send_command(
            InstanceIds=core_instance_ids,
            DocumentName="AWS-RunShellScript",
            Parameters={"commands": [command]},
            TimeoutSeconds=3600,
        )["Command"]["CommandId"]
        while True:
            # Verify the previous step succeeded before running the next step.
            cmd_result = ssm_client.list_commands(CommandId=command_id)["Commands"][0]
            if cmd_result["StatusDetails"] == "Success":
                print(f"Command succeeded.")
                break
            elif cmd_result["StatusDetails"] in ["Pending", "InProgress"]:
                print(f"Command status is {cmd_result['StatusDetails']}, waiting...")
                time.sleep(10)
            else:
                print(f"Command status is {cmd_result['StatusDetails']}, quitting.")
                raise RuntimeError(
                    f"Command {command} failed to run. "
                    f"Details: {cmd_result['StatusDetails']}"
```

```
                  )


def main():
    parser = argparse.ArgumentParser()
    parser.add_argument("cluster_id", help="The ID of the cluster.")
    parser.add_argument("script_path", help="The path to the script in Amazon S3.")
    args = parser.parse_args()

    emr_client = boto3.client("emr")
    ssm_client = boto3.client("ssm")

    install_libraries_on_core_nodes(
        args.cluster_id, args.script_path, emr_client, ssm_client
    )


if __name__ == "__main__":
    main()
```

JupyterHub 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 JupyterHub의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

JupyterHub 버전 정보

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-7.9.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
		client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.8.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.7.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-7.6.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.5.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-7.4.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.3.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-7.2.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.2	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-7.1.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-7.0.0	1.5.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.15.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.14.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.13.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.12.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.11.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.11.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.10.1	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.10.0	1.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.9.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.9.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.8.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.8.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.7.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.36.1	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.36.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.6.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.35.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.5.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.4.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.3.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.3.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.2.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.2.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.1.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.1.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-6.0.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-6.0.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.34.0	1.4.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.33.1	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.33.0	1.2.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.32.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.32.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.31.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.31.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.30.2	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.30.1	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.30.0	1.1.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.29.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.28.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.28.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.27.1	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.27.0	1.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.26.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.25.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.24.1	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.24.0	0.9.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.23.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.23.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.22.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.2	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.21.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.21.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.20.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.20.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.19.1	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.19.0	0.9.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.18.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.18.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.17.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.17.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.17.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.16.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.16.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.15.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.15.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.2	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Amazon EMR 릴리스 레이블	JupyterHub 버전	JupyterHub와 함께 설치된 구성 요소
emr-5.14.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub
emr-5.14.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, jupyterhub

Apache Livy

Livy는 REST 인터페이스를 통해 Spark를 실행하는 EMR 클러스터와 상호 작용할 수 있습니다. REST 인터페이스 또는 RPC 클라이언트 라이브러리를 사용하면 Spark 작업 또는 Spark 코드 조각을 전송하고 결과를 동기적 또는 비동기적으로 검색할 수 있으며 Spark 컨텍스트를 관리할 수 있습니다. 자세한 내용은 Apache Livy 웹 사이트를 참조하세요. Livy는 Amazon EMR 릴리스 버전 5.9.0 이상에 포함되어 있습니다.

Livy 웹 인터페이스에 액세스하려면 마스터 노드와 프록시 연결에 대한 SSH 터널을 설정하십시오. 자세한 내용은 EMR 클러스터에 호스팅된 웹 인터페이스 보기를 참조하세요.

다음 테이블에는 Amazon EMR이 Livy를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Livy의 버전이 나열되어 있습니다.

이 릴리스에서 Livy와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Livy 버전 정보

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-7.9.0	Livy 0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

다음 테이블에는 Amazon EMR이 Livy를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Livy의 버전이 나열되어 있습니다.

이 릴리스에서 Livy와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Livy 버전 정보

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.15.0	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

다음 테이블에는 Amazon EMR이 Livy를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Livy의 버전이 나열되어 있습니다.

이 릴리스에서 Livy와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Livy 버전 정보

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.36.2	Livy 0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server,

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
		spark-on-yarn, spark-yarn-slave, livy-server, nginx

주제

- [Apache Livy에서 HTTPS 활성화](#)
- [Livy 릴리스 기록](#)

Apache Livy에서 HTTPS 활성화

이 주제는 Amazon EMR 7.3.0 또는 이전 릴리스를 실행하는 경우 관련이 있습니다. 릴리스 7.4.0부터 HTTPS는 기본적으로 Apache Livy에서 활성화됩니다.

1. 전송 암호화가 활성화된 상태로 Amazon EMR 클러스터를 프로비저닝합니다. 암호화에 대해 자세히 알아보려면 [저장 데이터 및 전송 중 데이터 암호화](#)를 참조하세요.
2. 다음 콘텐츠를 가진 `livy_ssl.sh`이라는 파일을 생성합니다:

```
#!/bin/bash

KEYSTORE_FILE=`awk '/ssl.server.keystore.location/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEYSTORE_PASS=`awk '/ssl.server.keystore.password/{getline; print}' /etc/hadoop/
conf/ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`
KEY_PASS=`awk '/ssl.server.keystore.keypassword/{getline; print}' /etc/hadoop/conf/
ssl-server.xml | sed -e 's/<[^>]*>//g' | tr -d ' \t\n\r\f'`

echo "livy.keystore $KEYSTORE_FILE
livy.keystore.password $KEYSTORE_PASS
livy.key-password $KEY_PASS" | sudo tee -a /etc/livy/conf/livy.conf >/dev/null

sudo systemctl restart livy-server.service
```

3. 다음 스크립트를 Amazon EMR 단계로 실행합니다. 이 스크립트는 SSL을 활성화하도록 `/etc/livy/conf/livy.conf`를 수정합니다.

```
--steps '[{"Args":["s3://amzn-s3-demo-bucket/
livy_ssl.sh"],"Type":"CUSTOM_JAR","ActionOnFailure":"CONTINUE","Jar":"s3://
```

```
us-east-1.elasticmapreduce/libs/script-runner/script-
runner.jar","Properties":"","Name":"Custom JAR"}]'
```

4. 변경 사항을 적용하도록 Apache Livy 서비스를 다시 시작합니다. Apache Livy를 다시 시작하려면 프로세스 중지 및 다시 시작을 참조하세요.

5. 이제 클라이언트가 HTTPS를 사용하여 통신할 수 있는지 테스트합니다. 예를 들어 작업을 제출하려면 다음 코드를 실행합니다.

```
curl -k -X POST --data '{"file": "local:///usr/lib/spark/examples/jars/spark-
examples.jar",
"className": "org.apache.spark.examples.SparkPi"}' \
-H "Content-Type: application/json" \
https://EMR_Master_Node_Host:8998/batches
```

HTTPS를 활성화한 경우 Livy는 명령이 수락되었고 배치 작업이 제출되었음을 알리는 응답을 보냅니다.

```
{"id":1,"name":null,"owner":null,"proxyUser":null,"state":"starting","appId":null,"appInfo":
{"driverLogUrl":null,"sparkUiUrl":null},"log":["stdout: ","\nstderr: ","\nYARN
 Diagnostics: "]}
```

Livy 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Livy의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Livy 버전 정보

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-7.9.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
		yarn-resourcemanager, hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.8.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.7.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server , hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-7.6.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.5.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-7.4.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.3.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-7.2.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.2	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-7.1.0	0.8.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-7.0.0	0.7.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.15.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.14.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.13.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.12.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.11.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.11.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.10.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.10.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.9.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.9.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.8.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.8.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.7.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.36.1	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.36.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.6.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.35.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.5.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.4.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.3.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.3.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.2.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.2.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.1.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.1.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-6.0.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-6.0.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.34.0	0.7.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.33.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.33.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.32.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.32.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.31.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.31.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.30.2	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.30.1	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.30.0	0.7.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.29.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.28.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.28.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.27.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.27.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.26.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.25.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.24.1	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.24.0	0.6.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.23.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.23.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.22.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.21.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.21.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.20.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.20.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.19.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.19.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.18.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server, nginx
emr-5.18.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server, nginx

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.17.2	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.17.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.17.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.16.1	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.16.0	0.5.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.15.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.15.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.14.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.14.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.13.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.13.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.12.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.12.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.12.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-librar y, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-serv er, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.11.4	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.3	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.11.2	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.11.1	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.11.0	0.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.10.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Amazon EMR 릴리스 레이블	Livy 버전	Livy와 함께 설치된 구성 요소
emr-5.10.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.9.1	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server
emr-5.9.0	0.4.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, livy-server

Apache MXNet

> **Note**
>
> MXNet의 수명이 종료되었으며 프로젝트가 보관됩니다. Amazon EMR 7.2.0은 MXNet을 포함하는 마지막 릴리스입니다.

Apache MXNet는 신경망 및 기타 딥 러닝 애플리케이션을 구축하기 위해 설계된 가속화 라이브러리입니다. MXNet은 공통 워크플로우를 자동화하고 수치 계산을 최적화합니다. MXNet을 사용하면 선형 대수학 작업 같은 하위 수준 계산을 실행하는 데 초점을 맞출 필요 없이 신경망 아키텍처를 설계할 수 있습니다. MXNet은 Amazon EMR 릴리스 버전 5.10.0 이상에 포함되어 있습니다.

자세한 내용은 Apache MXNet 웹 사이트를 참조하십시오.

다음 테이블에는 Amazon EMR이 MXNet을 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 MXNet의 버전이 나열되어 있습니다.

MXNet과 함께 설치된 구성 요소의 버전은 릴리스 7.2.0 구성 요소 버전을 참조하세요.

emr-7.2.0용 MXNet 버전 정보

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-7.2.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

다음 테이블에는 Amazon EMR이 MXNet을 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 MXNet의 버전이 나열되어 있습니다.

이 릴리스에서 MXNet과 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 MXNet 버전 정보

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.15.0	MXNet 1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

다음 테이블에는 Amazon EMR이 MXNet을 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 MXNet의 버전이 나열되어 있습니다.

이 릴리스에서 MXNet과 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 MXNet 버전 정보

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.36.2	MXNet 1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

MXNet 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 MXNet의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

MXNet 버전 정보

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-7.2.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.2	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.1.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
		nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-7.0.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.15.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.14.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.13.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.12.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.11.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.11.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.10.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.10.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.9.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.9.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.8.1	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.8.0	1.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.7.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.36.1	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.36.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.6.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.35.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.5.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.4.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.3.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.3.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.2.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.2.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.1.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.1.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-6.0.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-6.0.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.34.0	1.8.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.33.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.33.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.32.1	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.32.0	1.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.31.1	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.31.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.2	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.30.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.30.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.29.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.28.1	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.28.0	1.5.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.27.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.27.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.26.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.25.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.24.1	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.24.0	1.4.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.23.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.23.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.22.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.2	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.21.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.21.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.20.1	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.20.0	1.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.19.1	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.19.0	1.3.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.18.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.18.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.2	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.17.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.17.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.16.1	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.16.0	1.2.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.15.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.15.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.14.2	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.1	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv
emr-5.14.0	1.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet, opencv

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.13.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.13.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.3	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.12.2	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.1	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.12.0	1.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.11.4	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.3	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.2	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.11.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.11.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet
emr-5.10.1	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Amazon EMR 릴리스 레이블	MXNet 버전	MXNet과 함께 설치된 구성 요소
emr-5.10.0	0.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mxnet

Apache Oozie

> **Note**
>
> Oozie 사용 중단 - OSS Oozie가 Attic 모드로 전환되었으므로 예정된 EMR 릴리스 중 하나에서 제거됩니다. 고객은 Oozie 대신 MWAA를 사용하는 것이 좋습니다. GitHub의 유틸리티 oozie-to-airflow-emr를 참조하여 기존 Oozie 워크플로를 MWAA로 마이그레이션할 수 있습니다.

Apache Oozie Workflow Scheduler를 사용하여 하둡 작업을 관리하고 조정합니다. 자세한 내용은 https://attic.apache.org/projects/oozie.html 참조하십시오.

Oozie 기본 웹 인터페이스는 Amazon EMR에서 지원되지 않습니다. Oozie에 대한 프런트 엔드 인스턴스를 사용하려면 Hue Oozie 애플리케이션을 사용해 보십시오. 자세한 내용은 Hue 단원을 참조하십시오. Oozie는 Amazon EMR 릴리스 버전 5.0.0 이상에 포함되어 있습니다. 이전 릴리스에서 Oozie는 샌드박스 애플리케이션으로 포함되어 있습니다. 자세한 내용은 Amazon EMR 4.x 릴리스 버전 단원을 참조하십시오.

생성 날짜가 2018-08-11인 Amazon Linux AMI를 기반으로 사용자 지정 Amazon Linux AMI를 사용하면 Oozie 서버가 시작에 실패합니다. Oozie를 사용하는 경우 생성 날짜가 다른 Amazon Linux AMI ID를 기반으로 사용자 지정 AMI를 생성하십시오. 다음 AWS CLI 명령을 사용하여 릴리스 날짜와 함께 2018.03 버전의 모든 HVM Amazon Linux AMIs에 대한 이미지 IDs 목록을 반환할 수 있으므로 적절한 Amazon Linux AMI를 기본으로 선택할 수 있습니다. MyRegion을 사용자의 리전 식별자(예: us-west-2)로 바꿉니다.

```
aws ec2 --region MyRegion describe-images --owner amazon --query 'Images[?
Name!=`null`]|[?starts_with(Name, `amzn-ami-hvm-2018.03`) == `true`].
[CreationDate,ImageId,Name]' --output text | sort -rk1
```

다음 테이블에는 Amazon EMR이 Oozie를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Oozie의 버전이 나열되어 있습니다.

이 릴리스에서 Oozie와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Oozie 버전 정보

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-7.9.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

다음 테이블에는 Amazon EMR이 Oozie를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Oozie의 버전이 나열되어 있습니다.

이 릴리스에서 Oozie와 함께 설치된 구성 요소의 버전은 [릴리스 6.15.0 구성 요소 버전]을 참조하세요.

emr-6.15.0용 Oozie 버전 정보

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.15.0	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resour

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
		cemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

다음 테이블에는 Amazon EMR이 Oozie를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Oozie의 버전이 나열되어 있습니다.

이 릴리스에서 Oozie와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Oozie 버전 정보

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.36.2	Oozie 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

주제

- Amazon RDS에서 원격 데이터베이스와 함께 Oozie 사용
- Oozie용 자바 버전 구성
- Oozie 릴리스 기록

Amazon RDS에서 원격 데이터베이스와 함께 Oozie 사용

기본적으로 Oozie 사용자 정보 및 쿼리 기록은 마스터 노드의 로컬 MySQL 데이터베이스에 저장됩니다. 또는 Amazon S3에 저장된 구성과 Amazon Relational Database Service(RDS)의 MySQL 데이터베이스를 사용하여 하나 이상의 Oozie 지원 클러스터를 생성할 수 있습니다. 이렇게 하면 Amazon EMR 클러스터의 실행 상태를 유지하지 않고도 Oozie에서 생성된 사용자 정보 및 쿼리 기록을 유지할 수 있습니다. Amazon S3 서버 측 암호화를 사용하여 구성 파일을 저장하는 것이 좋습니다.

먼저 Oozie용 원격 데이터베이스를 생성합니다.

외부 MySQL 데이터베이스를 생성하려면

1. https://console.aws.amazon.com/rds/에서 Amazon RDS 콘솔을 엽니다.
2. DB 인스턴스 시작을 선택합니다.
3. MySQL을 선택한 후 [Select]를 선택합니다.
4. 기본으로 선택된 Multi-AZ Deployment and Provisioned IOPS Storage(다중 AZ 배포 및 프로비저닝된 IOPS 스토리지)를 그대로 두고 다음을 선택합니다.
5. 인스턴스 사양에서 기본값을 그대로 사용하고, 설정을 지정하고, [Next]를 선택합니다.
6. [Configure Advanced Settings] 페이지에서 올바른 보안 그룹과 데이터베이스 이름을 선택합니다. 사용하는 보안 그룹은 클러스터의 마스터 노드에서 적어도 포트 3306에 대한 인바운드 TCP 액세스를 허용해야 합니다. 이때 클러스터를 생성하지 않은 경우 모든 호스트를 포트 3306에 연결하도록 허용하고 클러스터를 시작한 후 보안 그룹을 조정할 수 있습니다. [Launch DB Instance]를 선택합니다.
7. RDS 대시보드에서 인스턴스를 선택하고 방금 생성한 인스턴스를 선택합니다. 데이터베이스를 사용할 수 있는 경우 dbname, username, password 및 RDS instance hostname을 기록해 둡니다. 클러스터를 생성하고 구성할 때 이 정보를 사용합니다.

를 사용하여 클러스터를 시작할 때 Oozie에 대한 외부 MySQL 데이터베이스를 지정하려면 AWS CLI

AWS CLI를 사용하여 클러스터 시작 시 Oozie용 외부 MYSQL 데이터베이스를 지정하려는 경우 구성 객체로 `oozie-site`를 구성하기 위해 RDS 인스턴스를 생성할 때 적어둔 정보를 사용합니다.

> **Note**
>
> 동일한 외부 데이터베이스를 사용하는 여러 클러스터를 생성할 수 있지만, 각 클러스터는 쿼리 기록 및 사용자 정보를 공유합니다.

- 를 사용하여 Oozie가 설치된 클러스터를 AWS CLI생성하고, 생성한 외부 데이터베이스를 사용하고, 데이터베이스 속성을 지정하는 Oozie에 대한 구성 분류가 있는 구성 파일을 참조합니다. 다음은 Oozie가 설치된 클러스터를 생성하고 데이터베이스 구성을 지정하는 Amazon S3의 구성 파일 (`myConfig.json`)을 참조하는 예제입니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Oozie
 Name=Spark Name=Hive \
--instance-type m5.xlarge --instance-count 3 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/
myConfig.json --use-default-roles
```

다음은 `myConfig.json` 파일 콘텐츠 예제입니다. *JDBC URL*, *### ##* 및 *##*를 RDS 인스턴스의 JDBC URL, 사용자 이름 및 암호로 바꿉니다.

> **Important**
>
> JDBC URL에는 접미사로 데이터베이스 이름이 포함되어야 합니다. 예: jdbc:mysql://oozie-external-db.xxxxxxxxxx.us-east-1.rds.amazonaws.com:3306/dbname.

```
[{
  "Classification": "oozie-site",
    "Properties": {
        "oozie.service.JPAService.jdbc.driver": "org.mariadb.jdbc.Driver",
        "oozie.service.JPAService.jdbc.url": "JDBC URL",

        "oozie.service.JPAService.jdbc.username": "username",
        "oozie.service.JPAService.jdbc.password": "password"
    },
    "Configurations": []
}]
```

Oozie용 자바 버전 구성

Oozie는 여러 Java 가상 머신(JVM) 프로세스를 실행합니다. 이 페이지에서는 각 프로세스에 대해 Java 버전을 구성하는 방법을 설명합니다.

- Oozie 서버: `EmbeddedOozieServer`의 Java 버전을 업데이트하도록 `oozie-env` 분류에서 `JAVA_HOME`을 설정합니다.
- Oozie Launcher AM: Oozie Launcher AM은 Hadoop 및 Hive와 같은 적절한 애플리케이션 클라이언트 라이브러리를 간접 호출하는 단일 매퍼 MR 작업입니다. 달리 구성하지 않는 한, Oozie Launcher AM의 런타임 버전은 EMR 클러스터의 Hadoop용 Java 런타임과 동일합니다. Oozie Launcher AM용 Java 런타임을 구성하려면 작업의 `workflow.xml`에서 다음 속성을 설정합니다.

```
<property>
     <name>mapred.child.env</name>
     <value>JAVA_HOME=/path/to/JAVA_HOME</value>
 </property>
```

  이 속성을 사용하면 Oozie 작업의 Oozie Launcher AM이 Hadoop에 설정된 Java 버전이 아닌 사용자가 지정하는 Java 버전에서 실행됩니다.
- 애플리케이션 클라이언트 실행 파일: Oozie Launcher AM은 기본적으로 애플리케이션 클라이언트를 간접 호출하므로 클라이언트 실행 파일의 Java 런타임은 Oozie Launcher AM과 동일합니다.
- Oozie 작업으로 시작된 애플리케이션: 달리 지정하지 않는 한 Oozie 작업으로 실행되는 실제 애플리케이션 JVM의 런타임 버전은 EMR 클러스터의 Hadoop용 Java 런타임과 동일합니다. Oozie 작업(Spark 또는 Hive 작업)에서 애플리케이션을 시작하는 데 사용되는 Oozie 워크플로 작업의 유형에 따라 Oozie 작업의 `workflow.xml`에서 실제 애플리케이션 JVM의 기본 Java 런타임을 업데이트할 수 있습니다.

Oozie 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Oozie의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Oozie 버전 정보

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-7.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-7.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-7.2.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.36.2	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-7.1.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-7.0.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.15.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.14.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.13.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.12.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.11.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.11.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.10.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.10.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn, tez-on-worker
emr-6.9.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.8.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.36.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.36.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.6.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.35.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.5.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.4.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.3.1	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.3.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.2.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.2.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.1.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.1.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-6.0.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-6.0.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.34.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.33.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.33.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.32.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.32.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.31.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.31.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.30.2	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.30.1	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.30.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.29.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.28.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.28.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.27.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.27.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.26.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.25.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.24.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.24.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.23.1	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.23.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.22.0	5.1.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.21.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.21.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.21.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.20.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.20.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.19.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.19.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.18.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.18.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.17.2	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.17.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.17.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.16.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.16.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.15.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.15.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.14.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.14.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.14.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.13.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.13.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.12.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.12.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.12.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.11.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.11.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.11.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.10.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.10.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.9.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.9.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.8.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.8.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.8.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.7.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.7.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.6.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.6.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.5.4	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.3	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.5.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.5.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.5.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.4.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.4.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.3.2	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.3.1	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.3.0	4.3.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.2.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.2.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.2.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.1.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.1.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.0.3	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.2	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Oozie 버전	Oozie와 함께 설치된 구성 요소
emr-5.0.1	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn
emr-5.0.0	4.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, oozie-client, oozie-server, tez-on-yarn

버전별 Oozie 릴리스 정보

주제

- Amazon EMR 7.6.0 - Oozie 변경 사항
- Amazon EMR 7.5.0 - Oozie 변경 사항
- Amazon EMR 7.4.0 - Oozie 변경 사항

Amazon EMR 7.6.0 - Oozie 변경 사항

7.6.0 릴리스에는 변경 사항이 없습니다.

Amazon EMR 7.5.0 - Oozie 변경 사항

유형	설명
업그레이드	새 EMR 릴리스에서는 Sqoop을 사용할 수 없으므로 Sqoop 작업 오케스트레이션에 대한 지원 제거

Amazon EMR 7.4.0 - Oozie 변경 사항

Oozie 변경 사항: 변경 사항 없음

Amazon EMR 7.3.0 - Oozie 변경 사항

유형	설명
업그레이드	guice를 4.0으로 업그레이드합니다.
기본 동작 변경	`oozie-site.xml` 에서 두 속성을 정의합니다. • `oozie.service.SparkConfigurationService.spark.configurations=*=/etc/spark/conf/` • `oozie.service.SparkConfigurationService.spark.configurations.blacklist=spark.yarn.dist.files`

유형	설명
	이러한 속성을 사용하면 Oozie를 Spark와 함께 설치할 때 Oozie에서 기본 Amazon EMR Spark 구성을 선택할 수 있습니다.

Amazon EMR 7.1.0 - Oozie 변경 사항

유형	설명
업그레이드	Netty를 4.1.100.Final로 업그레이드합니다.
업그레이드	Jetty를 9.4.53.v20231009로 업그레이드합니다.

Amazon EMR 7.0.0 - Oozie 변경 사항

유형	설명
업그레이드	이제 Oozie 클라이언트 및 서버는 기본적으로 Java 17을 사용합니다. 자세한 내용은 EMR 7.0.0 릴리스 가이드를 참조하세요.

Apache Phoenix

Apache Phoenix는 OLTP 및 작동 분석에 사용되며, 표준 SQL 쿼리와 JDBC API를 Apache HBase 보조 저장소와 함께 작동할 수 있도록 합니다. 자세한 내용은 Phoenix in 15 minutes or less(15분 이내의 Phoenix)를 참조하십시오. Phoenix는 Amazon EMR 릴리스 버전 4.7.0 이상에 포함되어 있습니다.

이전 버전의 Amazon EMR에서 Amazon EMR 릴리스 버전 5.4.0 이상으로 업그레이드하고 보조 인덱싱을 사용하는 경우 Apache Phoenix 설명서에 설명된 대로 로컬 인덱스를 업그레이드합니다. Amazon EMR은 `hbase-site` 분류에서 필수 구성을 제거하지만 인덱스를 다시 채워야 합니다. 인덱스의 온라인 및 오프라인 업그레이드가 지원됩니다. 온라인 업그레이드는 기본값이므로, 버전 4.8.0 이상인 Phoenix 클라이언트에서 초기화하는 동안 인덱스가 다시 채워집니다. 오프라인 업그레이드를 지정하려면 `phoenix.client.localIndexUpgrade` 분류에서 `phoenix-site` 구성을 false로 설정한 다음 SSH를 마스터 노드로 설정하여 `psql [zookeeper] -1`을 실행합니다.

다음 테이블에는 Amazon EMR이 Phoenix를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Phoenix의 버전이 나열되어 있습니다.

이 릴리스에서 Phoenix와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Phoenix 버전 정보

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.9.0	Phoenix 5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
		connectors, phoenix-query-server, zookeeper-client, zookeeper-server

다음 테이블에는 Amazon EMR이 Phoenix를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Phoenix의 버전이 나열되어 있습니다.

이 릴리스에서 Phoenix와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Phoenix 버전 정보

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.15.0	Phoenix 5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

다음 테이블에는 Amazon EMR이 Phoenix를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Phoenix의 버전이 나열되어 있습니다.

이 릴리스에서 Phoenix와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Phoenix 버전 정보

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.36.2	Phoenix 4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

주제

- Phoenix를 포함하는 클러스터 생성
- Phoenix 클라이언트
- Phoenix 릴리스 기록

Phoenix를 포함하는 클러스터 생성

콘솔이나 AWS CLI를 사용하여 클러스터를 생성할 때 애플리케이션을 선택하여 Phoenix를 설치할 수 있습니다. 다음 절차 및 예제는 Phoenix 및 HBase를 포함하는 클러스터를 생성하는 방법을 보여 줍니다. 고급 옵션을 비롯해 콘솔을 사용하여 클러스터를 생성하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 계획 및 구성을 참조하세요.

콘솔에서 클러스터를 생성하기 위해 빠른 옵션을 사용하여 Phoenix를 설치한 상태로 클러스터를 시작하려면

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 빠른 생성을 사용하려면 클러스터 생성을 선택합니다.
3. 소프트웨어 구성에서 애플리케이션에 해당하는 최신 릴리스를 선택합니다. Phoenix는 Amazon Release Version emr-4.7.0 이상을 선택한 경우에만 옵션으로 표시됩니다.
4. 애플리케이션에서 두 번째 옵션인 HBase: HBase ***ver*** with Ganglia ***ver***, Hadoop ***ver***, Hive ***ver***, Hue ***ver***, Phoenix ***ver*** 및 ZooKeeper ***ver***를 선택합니다.
5. 필요에 따라 다른 옵션을 선택한 다음 클러스터 생성을 선택합니다.

> **Note**
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

다음 예제에서는 기본 구성 설정을 사용하여 Phoenix가 설치된 상태로 클러스터를 시작합니다.

를 사용하여 Phoenix 및 HBase로 클러스터를 시작하려면 AWS CLI

- 다음 명령을 사용하여 클러스터를 생성합니다.

```
aws emr create-cluster --name "Cluster with Phoenix" --release-label emr-7.9.0 \
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Phoenix 구성 사용자 지정

클러스터 생성 시 `hbase-site` 구성 분류를 사용하여 `hbase-site.xml`에서 값을 설정하는 방식으로 Phoenix를 구성합니다.

자세한 내용은 Phoenix 설명서에서 Configuration and Tuning을 참조하세요.

다음 예제에서는 Amazon S3에 저장된 JSON 파일을 사용하여 `phoenix.schema.dropMetaData` 속성에 대해 `false` 값을 지정하는 방법을 보여줍니다. 단일 분류에 대해 여러 속성을 지정할 수 있습니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오. 그런 다음 `create-cluster` 명령이 JSON 파일을 `--configurations` 파라미터로 참조합니다.

/amzn-s3-demo-bucket/myfolder/myconfig.json에 저장된 JSON 파일의 콘텐츠는 다음과 같습니다.

```
[
    {
      "Classification": "hbase-site",
      "Properties": {
        "phoenix.schema.dropMetaData": "false"
      }
    }
  ]
```

JSON 파일을 참조하는 `create cluster` 명령이 다음 예제에 나와 있습니다.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Phoenix \
Name=HBase --instance-type m5.xlarge --instance-count 2 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myconfig.json
```

Note

Phoenix 구성 분류에 대한 재구성 요청은 Amazon EMR 버전 5.23.0 이상에서만 지원되고 Amazon EMR 버전 5.21.0 또는 5.22.0에서는 지원되지 않습니다. 자세한 내용은 실행 중인 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하세요.

Phoenix 클라이언트

모든 종속 조건으로 빌드된 JDBC 클라이언트를 사용하거나 Phoenix Query Server를 사용하고 클러스터의 마스터 노드에서만 실행할 수 있는(예를 들면 SQL 클라이언트, 단계 명령줄, SSH 포트 전달 등을

통해) "씬(thin) 클라이언트"를 사용하여 Phoenix에 연결할 수 있습니다. "팻(fat)" JDBC 클라이언트를 사용하는 경우 HBase 서비스에 직접 연결되므로 여전히 클러스터의 모든 노드에 액세스할 수 있어야 합니다. "씬" Phoenix 클라이언트는 기본 포트 8765에서만 Phoenix Query Server에 액세스할 수 있으면 됩니다. Phoenix 안에는 이러한 클라이언트를 사용하는 여러 스크립트가 있습니다.

Amazon EMR 단계를 사용하여 Phoenix를 통해 쿼리

다음 절차에서는 HBase의 스냅샷을 복원한 후 해당 데이터를 사용하여 Phoenix 쿼리를 실행합니다. 자신의 요구에 맞게 Phoenix 클라이언트를 활용하는 새 스크립트를 생성하거나 이 예제를 확장할 수 있습니다.

1. 다음 명령을 사용하여 Phoenix가 설치된 상태로 클러스터를 생성합니다.

```
aws emr create-cluster --name "Cluster with Phoenix" --log-uri s3://amzn-s3-demo-bucket/myLogFolder --release-label emr-7.9.0 \
--applications Name=Phoenix Name=HBase --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

2. 다음 파일을 생성하여 Amazon S3에 업로드합니다.

   copySnapshot.sh

```
sudo su hbase -s /bin/sh -c 'hbase snapshot export \
 -D hbase.rootdir=s3://us-east-1.elasticmapreduce.samples/hbase-demo-customer-data/snapshot/ \
-snapshot customer_snapshot1 \
-copy-to hdfs://masterDNSName:8020/user/hbase \
-mappers 2 -chuser hbase -chmod 700'
```

   runQuery.sh

```
aws s3 cp s3://amzn-s3-demo-bucket/phoenixQuery.sql /home/hadoop/
/usr/lib/phoenix/bin/sqlline-thin.py http://localhost:8765 /home/hadoop/phoenixQuery.sql
```

   phoenixQuery.sql

> **Note**
>
> Amazon EMR 버전 5.26.0 이상을 사용하는 경우 다음 예제에 `COLUMN_ENCODED_BYTES=0`을 포함하기만 하면 됩니다.

```
CREATE VIEW "customer" (
pk VARCHAR PRIMARY KEY,
"address"."state" VARCHAR,
"address"."street" VARCHAR,
"address"."city" VARCHAR,
"address"."zip" VARCHAR,
"cc"."number" VARCHAR,
"cc"."expire" VARCHAR,
"cc"."type" VARCHAR,
"contact"."phone" VARCHAR)
COLUMN_ENCODED_BYTES=0;

CREATE INDEX my_index ON "customer" ("customer"."state") INCLUDE("PK",
 "customer"."city", "customer"."expire", "customer"."type");

SELECT "customer"."type" AS credit_card_type, count(*) AS num_customers FROM
 "customer" WHERE "customer"."state" = 'CA' GROUP BY "customer"."type";
```

AWS CLI 를 사용하여 파일을 S3 버킷에 제출합니다.

```
aws s3 cp copySnapshot.sh s3://amzn-s3-demo-bucket/
aws s3 cp runQuery.sh s3://amzn-s3-demo-bucket/
aws s3 cp phoenixQuery.sql s3://amzn-s3-demo-bucket/
```

3. 1단계에서 생성한 클러스터로 제출된 다음 단계를 사용하여 테이블을 생성합니다.

   createTable.json

```
[
  {
    "Name": "Create HBase Table",
    "Args": ["bash", "-c", "echo $'create \"customer\",\"address\",\"cc\",\"contact
\"' | hbase shell"],
    "Jar": "command-runner.jar",
```

```
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./createTable.json
```

4. `script-runner.jar`를 사용하여 S3 버킷으로 이전에 업로드한 `copySnapshot.sh` 스크립트를 실행합니다.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="HBase Copy Snapshot",ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/copySnapshot.sh"]
```

   이렇게 하면 MapReduce 작업이 스냅샷 데이터를 클러스터 HDFS로 복사합니다.

5. 다음 단계를 사용하여 클러스터로 복사한 스냅샷을 복원합니다.

   restoreSnapshot.json

```
[
  {
    "Name": "restore",
    "Args": ["bash", "-c", "echo $'disable \"customer\"; restore_snapshot
 \"customer_snapshot1\"; enable \"customer\"' | hbase shell"],
    "Jar": "command-runner.jar",
    "ActionOnFailure": "CONTINUE",
    "Type": "CUSTOM_JAR"
  }
]
```

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps file://./restoreSnapshot.json
```

6. `script-runner.jar`를 사용하여 S3 버킷으로 이전에 업로드한 `runQuery.sh` 스크립트를 실행합니다.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Phoenix Run Query",ActionOnFailure=CONTINUE,\
```

```
Jar=s3://region.elasticmapreduce/libs/script-runner/script-
runner.jar,Args=["s3://amzn-s3-demo-bucket/runQuery.sh"]
```

쿼리가 실행되고 결과가 단계의 `stdout`로 반환됩니다. 이 단계를 완료하는 데 몇 분이 걸릴 수도 있습니다.

7. 1단계에서 클러스터를 생성할 때 사용한 로그 URI에서 해당 단계의 `stdout` 결과를 검사합니다. 결과는 다음과 같아야 합니다.

```
+------------------------------------------+-----------------------------------+
|            CREDIT_CARD_TYPE              |           NUM_CUSTOMERS           |
+------------------------------------------+-----------------------------------+
| american_express                         | 5728                              |
| dankort                                  | 5782                              |
| diners_club                              | 5795                              |
| discover                                 | 5715                              |
| forbrugsforeningen                       | 5691                              |
| jcb                                      | 5762                              |
| laser                                    | 5769                              |
| maestro                                  | 5816                              |
| mastercard                               | 5697                              |
| solo                                     | 5586                              |
| switch                                   | 5781                              |
| visa                                     | 5659                              |
+------------------------------------------+-----------------------------------+
```

Phoenix 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Phoenix의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Phoenix 버전 정보

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.9.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
		cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server
emr-7.8.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.7.0	5.2.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.6.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.5.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.4.0	5.2.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.3.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.2.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.36.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.1.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-7.0.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.15.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.14.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.13.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.12.0	5.1.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.11.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.11.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.10.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.10.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.9.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.9.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.8.1	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.8.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.7.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.36.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.36.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.6.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, hbase-operator-tools, phoenix-library, phoenix-connectors, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.35.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.5.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.4.0	5.1.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.3.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.3.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.2.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.2.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.1.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.1.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-6.0.1	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-6.0.0	5.0.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.34.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.33.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.33.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.32.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.32.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.31.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.31.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.30.2	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.30.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.30.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.29.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.28.1	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.28.0	4.14.3	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.27.1	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.27.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.26.0	4.14.2	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.25.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.24.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.24.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.23.1	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.23.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.22.0	4.14.1	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.21.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.21.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.20.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.20.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.19.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.19.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.18.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.18.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.17.2	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.17.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.17.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.16.1	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.16.0	4.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.15.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.15.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.14.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.14.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.13.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.13.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.12.3	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.2	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.12.1	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.12.0	4.13.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.11.4	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.11.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.11.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.11.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.10.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.10.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.9.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.9.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.3	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.8.2	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.8.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.8.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.7.1	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.7.0	4.11.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.6.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.6.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.4	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.5.3	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.2	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.5.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.5.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.4.1	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.4.0	4.9.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.3.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.3.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.3.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.2.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.2.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.2.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.1.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.1.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.0.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-5.0.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-5.0.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.6	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.9.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.9.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.9.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.9.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.5	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.8.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.8.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.8.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.8.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.4	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.7.3	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.2	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	Phoenix 버전	Phoenix와 함께 설치된 구성 요소
emr-4.7.1	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server
emr-4.7.0	4.7.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-mapred, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hbase-hmaster, hbase-client, hbase-region-server, phoenix-library, phoenix-query-server, zookeeper-client, zookeeper-server

Apache Pig

Apache Pig는 하둡을 기반으로 실행되는 오픈 소스 Apache 라이브러리이며, Java와 같은 낮은 수준의 컴퓨터 언어로 복잡한 코드를 작성할 필요가 없이 대규모 데이터 세트를 변환하는 데 사용할 수 있는 스크립팅 언어를 제공합니다. 라이브러리는 Pig Latin이라는 언어로 작성된 SQL 유형 명령을 가져와서, DAG(방향성 비순환 그래프) 또는 MapReduce 프로그램 기반의 Tez 작업으로 변환합니다. Pig는 다양한 형식의 정형 및 비정형 데이터를 사용하여 작업합니다. Pig에 대한 자세한 내용은 http://pig.apache.org/를 참조하십시오.

대화형으로 또는 배치 모드에서 Pig 명령을 실행할 수 있습니다. 대화형으로 Pig를 사용하려면 마스터 노드에 대한 SSH 연결을 생성하고 Grunt 셸을 사용하여 명령을 제출합니다. 배치 모드로 Pig를 사용하려면 Pig 스크립트를 작성하고 해당 스크립트를 Amazon S3에 업로드한 다음 클러스터 단계로 제출합니다. 클러스터에 작업을 제출하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터에 작업 제출을 참조하세요.

Pig를 사용하여 Amazon S3의 HCatalog 테이블에 출력을 쓰는 경우 `mapred.output.direct.NativeS3FileSystem` 및 `mapred.output.direct.EmrFileSystem` 속성을 `false`로 설정하여 Amazon EMR 직접 쓰기를 비활성화합니다. 자세한 내용은 HCatalog 사용 단원을 참조하십시오. Pig 스크립트에서 `SET mapred.output.direct.NativeS3FileSystem false` 및 `SET mapred.output.direct.EmrFileSystem false` 명령을 사용할 수 있습니다.

다음 테이블에는 Amazon EMR이 Pig를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Pig의 버전이 나열되어 있습니다.

이 릴리스에서 Pig와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Pig 버전 정보

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-7.9.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

다음 테이블에는 Amazon EMR이 Pig를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Pig의 버전이 나열되어 있습니다.

이 릴리스에서 Pig와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Pig 버전 정보

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.15.0	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

다음 테이블에는 Amazon EMR이 Pig를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Pig의 버전이 나열되어 있습니다.

이 릴리스에서 Pig와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Pig 버전 정보

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.36.2	Pig 0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

주제

- Pig 작업 제출
- Pig에서 사용자 정의 함수 직접 호출
- Pig 릴리스 기록

Pig 작업 제출

이 섹션에서는 Amazon EMR 클러스터에 Pig 작업을 제출하는 방법을 보여줍니다. 다음 예제는 전송된 총 바이트 수, 상위 50개 IP 주소 목록, 상위 50개 외부 참조자 목록, Bing과 Google을 사용한 상위 50개 검색어 등이 포함된 보고서를 생성합니다. Pig 스크립트는 Amazon S3 버킷(`s3://elasticmapreduce/samples/pig-apache/do-reports2.pig`)에 있습니다. 입력 데이터는 Amazon S3 버킷(`s3://elasticmapreduce/samples/pig-apache/input`)에 있습니다. 출력은 Amazon S3 버킷에 저장됩니다.

Amazon EMR 콘솔을 사용하여 Pig 작업 제출

이 예제는 Amazon EMR 콘솔을 사용하여 Pig 단계를 클러스터에 추가하는 방법을 설명합니다.

Pig 단계를 제출하려면

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 생성을 선택하여 Pig가 설치된 클러스터를 생성합니다. 클러스터를 생성하는 방법에 대한 단계는 Amazon EMR 클러스터 계획 및 구성을 참조하세요.
3. SSH를 사용하여 프라이머리 노드에 연결에 설명된 단계에 따라 터미널을 열고 클러스터의 프라이머리 노드에 SSH로 연결합니다. 완료했으면 다음 단계를 실행합니다.

```
sudo mkdir -p /home/hadoop/lib/pig/
sudo aws s3 cp s3://elasticmapreduce/libs/pig/0.3/piggybank-0.3-amzn.jar /home/
hadoop/lib/pig/piggybank.jar
```

4. 콘솔에서 클러스터 목록을 클릭하고 생성한 클러스터의 이름을 선택합니다.
5. 단계 섹션으로 스크롤하여 섹션을 확장한 후 단계 추가를 선택합니다.
6. 단계 추가 대화 상자에서 다음을 수행합니다.
   - 단계 유형에서 Pig 프로그램을 선택합니다.
   - 이름에서 기본 이름(Pig 프로그램)을 수락하거나 새 이름을 입력합니다.
   - 스크립트 S3 위치에 Pig 스크립트의 위치를 입력합니다. 예를 들어 **s3://elasticmapreduce/samples/pig-apache/do-reports2.pig**입니다.
   - 입력 S3 위치에 입력 데이터의 위치를 입력합니다. 예를 들어 **s3://elasticmapreduce/samples/pig-apache/input**입니다.
   - 출력 S3 위치에서 Amazon S3 출력 버킷의 이름을 입력하거나 찾아봅니다.
   - 인수에서 필드를 비워 둡니다.
   - 실패 시 작업에서 기본 옵션(계속)을 그대로 사용합니다.
7. 추가를 선택합니다. 단계가 콘솔에 [Pending] 상태로 나타납니다.
8. 단계의 상태는 단계가 실행됨에 따라 대기 중에서 실행 중을 거쳐 완료됨으로 바뀝니다. 상태를 업데이트하려면 작업 열 위의 새로 고침 아이콘을 선택합니다. 단계가 완료되면 Pig 단계의 출력 파일이 있는지 Amazon S3 버킷을 확인합니다.

를 사용하여 Pig 작업 제출 AWS CLI

를 사용하여 Pig 단계를 제출하려면 AWS CLI

를 사용하여 클러스터를 시작할 때 `--applications` 파라미터를 AWS CLI사용하여 Pig를 설치합니다. Pig 단계를 제출하려면 `--steps` 파라미터를 사용합니다.

1. Pig가 설치된 클러스터를 시작하려면 다음 명령을 입력합니다. 이때 *mykey* 및 *amzn-s3-demo-bucket*을 각각 EC2 키 페어 및 Amazon S3 버킷으로 바꿉니다.

```
aws emr create-cluster \
--name "Test cluster" \
--log-uri s3://amzn-s3-demo-bucket/ \
--release-label emr-5.36.2 \
--applications Name=Pig \
--use-default-roles \
--ec2-attributes KeyName=myKey \
--instance-type m5.xlarge \
--instance-count 3
```

> Note
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`--instance-groups` 파라미터를 사용하지 않고 인스턴스 수를 지정하면 단일 마스터 노드가 시작되고 나머지 인스턴스는 코어 노드로 시작됩니다. 모든 노드에는 이 명령에 지정된 인스턴스 유형이 사용됩니다.

> Note
>
> EMR 서비스 역할과 EC2 인스턴스 프로파일을 아직 생성하지 않았다면 `aws emr create-default-roles` 하위 명령을 입력하기 전에 `create-cluster`를 입력하여 생성합니다.

2. Pig 단계를 제출하려면 다음 명령을 입력합니다. 이때 *myClusterId* 및 *amzn-s3-demo-bucket*을 각각 클러스터 ID 및 Amazon S3 버킷 이름으로 바꿉니다.

```
aws emr add-steps \
--cluster-id myClusterId \
--steps Type=PIG,Name="Pig Program",ActionOnFailure=CONTINUE,Args=[-f,s3://
elasticmapreduce/samples/pig-apache/do-reports2.pig,-p,INPUT=s3://elasticmapreduce/
samples/pig-apache/input,-p,OUTPUT=s3://amzn-s3-demo-bucket/pig-apache/output]
```

이 명령은 단계의 `State`를 확인하는 데 사용할 수 있는 단계 ID를 반환합니다.

3. `describe-step` 명령을 사용하여 단계 상태를 쿼리합니다.

```
aws emr describe-step --cluster-id myClusterId --step-id s-1XXXXXXXXXXA
```

   단계가 실행됨에 따라 단계의 `State`가 `PENDING`에서 `RUNNING`, 그리고 `COMPLETED`로 바뀝니다. 단계가 완료되면 Pig 단계의 출력 파일이 있는지 Amazon S3 버킷을 확인합니다.

에서 Amazon EMR 명령을 사용하는 방법에 대한 자세한 AWS CLI내용은 AWS CLI 명령 참조를 참조하세요.

Pig에서 사용자 정의 함수 직접 호출

Pig는 Pig 스크립트 내에서 사용자 정의 함수(UDF)를 호출할 수 있는 기능을 제공합니다. 이 기능을 사용하여 Pig 스크립트에서 사용할 사용자 지정 처리를 구현할 수 있습니다. 현재 지원되는 언어는 Java, Python/Jython 및 JavaScript(JavaScript 지원은 아직 실험적임)입니다.

다음 단원에서는 Pig 셸에서 또는 Pig 스크립트 내에서 호출할 수 있도록 Pig에 함수를 등록하는 방법을 설명합니다. Pig에서 UDF를 사용하는 방법에 자세한 내용은 해당 Pig 버전의 Pig 설명서를 참조하세요.

Pig에서 JAR 파일 호출

Pig 스크립트에서 `REGISTER` 명령을 사용하여 Pig에서 사용자 지정 JAR 파일을 사용할 수 있습니다. JAR 파일은 로컬 파일 시스템이거나 Amazon S3와 같은 원격 파일 시스템입니다. Pig 스크립트가 실행되면 Amazon EMR은 JAR 파일을 프라이머리 노드에 자동으로 다운로드한 다음 JAR 파일을 Hadoop 분산 캐시에 업로드합니다. 이 방식으로 JAR 파일은 클러스터의 모든 인스턴스에서 필요에 따라 자동으로 사용됩니다.

Pig에서 JAR 파일을 사용하려면

1. 사용자 지정 JAR 파일을 Amazon S3에 업로드합니다.
2. Pig 스크립트에서 `REGISTER` 명령을 사용하여 사용자 지정 JAR 파일의 Amazon S3에 있는 버킷을 지정합니다.

```
REGISTER s3://amzn-s3-demo-bucket/path/mycustomjar.jar;
```

Pig에서 Python 및 Jython 스크립트 직접 호출

Python 스크립트를 Pig에 등록한 다음 Pig 셸의 해당 스크립트 또는 Pig 스크립트에서 함수를 호출할 수 있습니다. `register` 키워드로 스크립트의 위치를 지정하여 이 작업을 수행합니다.

Pig는 Java로 작성되기 때문에 Jython 스크립트 엔진을 사용하여 Python 스크립트를 구문 분석합니다. Jython에 대한 자세한 내용은 http://www.jython.org/를 참조하십시오.

Pig에서 Python/Jython 스크립트를 호출하려면

1. Python 스크립트를 작성하고 스크립트를 Amazon S3의 한 위치에 업로드합니다. 이 위치는 Pig 클러스터를 생성하는 계정 또는 클러스터를 생성한 계정이 클러스터에 액세스할 수 있도록 권한이 설정된 계정과 동일한 계정이 소유한 버킷이어야 합니다. 이 예제에서 스크립트는 s3://amzn-s3-demo-bucket/pig/python에 업로드됩니다.
2. Pig 클러스터를 시작합니다. Grunt 셸에서 Pig에 액세스할 경우 대화형 클러스터를 실행합니다. 스크립트에서 Pig 명령을 실행할 경우 스크립팅된 Pig 클러스터를 시작합니다. 이 예제에서는 대화형 클러스터가 시작됩니다. Pig 클러스터를 생성하는 방법에 대한 자세한 내용은 Pig 작업 제출 단원을 참조하십시오.
3. 대화형 클러스터의 경우, SSH를 사용하여 마스터 노드에 연결한 후 Grunt 셸을 실행합니다. 자세한 내용은 프라이머리 노드에 대한 SSH 연결을 참조하세요.
4. 명령줄에 `pig`를 입력하여 Pig용 Grunt 셸을 실행합니다.

```
pig
```

5. 다음과 같이 Grunt 명령 프롬프트에서 `register` 키워드를 사용하여 Jython 라이브러리와 Python 스크립트를 Pig에 등록합니다. 이때 Amazon S3에서 스크립트의 위치를 지정합니다.

```
grunt> register 'lib/jython.jar';
grunt> register 's3://amzn-s3-demo-bucket/pig/python/myscript.py' using jython as
 myfunctions;
```

6. 입력 데이터를 로드합니다. 다음 예제는 Amazon S3 위치에서 입력을 로드합니다.

```
grunt> input = load 's3://amzn-s3-demo-bucket/input/data.txt' using TextLoader as
 (line:chararray);
```

7. 이제 `myfunctions`를 사용하여 함수를 참조하면 Pig 내에서 스크립트에 함수를 호출할 수 있습니다.

```
grunt> output=foreach input generate myfunctions.myfunction($1);
```

Pig 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Pig의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Pig 버전 정보

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-7.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
		timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-7.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-7.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-7.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.36.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-7.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-7.0.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn, tez-on-worker
emr-6.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.8.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.8.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.7.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.36.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.36.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.6.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.35.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.5.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.4.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.3.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.3.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.2.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.2.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-6.1.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-6.1.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.34.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.33.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.33.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.32.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.32.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.31.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.31.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.30.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.30.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.30.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.29.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.28.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.28.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.27.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.27.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.26.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.25.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.24.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.24.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.23.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.23.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.22.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.21.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.21.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.21.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.20.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.20.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.19.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.19.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.18.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.18.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.17.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.17.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.17.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.16.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.16.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.15.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.15.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.14.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.14.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.14.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.13.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.13.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.12.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.12.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.12.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.4	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.11.3	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.2	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.11.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.11.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.10.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.10.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.9.1	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.9.0	0.17.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.8.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.8.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.8.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.7.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.7.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.6.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.6.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.5.4	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.5.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.5.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.5.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.4.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.4.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.3.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.3.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.3.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.2.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.2.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.2.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.1.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.1.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.0.3	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.2	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-5.0.1	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn
emr-5.0.0	0.16.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, pig-client, tez-on-yarn

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.9.6	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.9.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.9.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.8.5	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.8.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.8.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.7.4	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.3	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.2	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.7.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.7.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.6.1	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.6.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.5.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.4.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.3.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.2.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client
emr-4.1.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Amazon EMR 릴리스 레이블	Pig 버전	Pig와 함께 설치된 구성 요소
emr-4.0.0	0.14.0	emrfs, emr-ddb, emr-goodies, emr-kinesis, emr-s3-dist-cp, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, pig-client

Presto

> **Important**
>
> 이전 버전의 Trino인 Presto는 Amazon EMR에서 계속 사용할 수 있습니다. 하지만 앞으로 Amazon EMR에서 사용할 Trino를 적극 권장합니다. 또한 Trino와 Presto는 동일한 클러스터에서 동시에 실행할 수 없습니다. 자세한 내용은 Trino를 참조하세요.

Presto는 여러 소스의 대규모 데이터 세트에 대한 대화형 분석 쿼리를 위해 설계된 빠른 SQL 쿼리 엔진입니다. 자세한 내용은 Presto 웹 사이트를 참조하세요. Presto는 Amazon EMR 릴리스 5.0.0 이상에 포함되어 있습니다. 이전 릴리스에는 Presto가 샌드박스 애플리케이션으로 포함되었습니다. 자세한 내용은 Amazon EMR 4.x 릴리스 버전 단원을 참조하십시오. Amazon EMR 릴리스 버전 6.1.0 이상은 Presto 외에도 Trino를 지원합니다. 자세한 내용은 PrestoDB 설치 단원을 참조하십시오.

다음 테이블에는 Amazon EMR이 Presto를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Presto의 버전이 나열되어 있습니다.

이 릴리스에서 Presto와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Presto 버전 정보

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-7.9.0	Presto 0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

다음 테이블에는 Amazon EMR이 Presto를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Presto의 버전이 나열되어 있습니다.

이 릴리스에서 Presto와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Presto 버전 정보

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.15.0	Presto 0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

다음 테이블에는 Amazon EMR이 Presto를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Presto의 버전이 나열되어 있습니다.

이 릴리스에서 Presto와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Presto 버전 정보

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.36.2	Presto 0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
		resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

주제

- AWS Glue 데이터 카탈로그에서 Presto 사용
- S3 Select Pushdown과 함께 Presto를 사용하여 성능 향상
- 데이터베이스 커넥터 추가
- Amazon EMR 기반 Presto를 통해 SSL 및 TLS 사용과 LDAPS 구성
- Presto 엄격한 모드 활성화
- Presto에서 스팟 인스턴스 손실 처리
- 단계적 서비스 해제 기능이 있는 자동 조정 사용
- Amazon EMR 기반 Presto에 대한 고려 사항
- Presto 릴리스 기록

AWS Glue 데이터 카탈로그에서 Presto 사용

Amazon EMR 릴리스 버전 5.10.0 이상을 사용하여 AWS Glue 데이터 카탈로그를 Presto의 기본 Hive 메타스토어로 지정할 수 있습니다. 영구 메타스토어가 필요하거나 여러 클러스터, 서비스, 애플리케이션 또는 AWS 계정에서 메타스토어를 공유해야 하는 경우에 이 구성을 사용하는 것이 좋습니다.

AWS Glue는 완전 관리형 추출, 변환 및 로드(ETL) 서비스로, 데이터를 분류하고 정리하고 보강하고 다양한 데이터 스토어 간에 안정적으로 이동할 수 있는 간단하고 비용 효율적입니다. Glue 데이터 카탈로그는 Amazon EMR뿐만 아니라 Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena 및 Apache Hive AWS 메타스토어와 호환되는 모든 애플리케이션과 통합하여 다양한 데이터 소스 및 데이터 형식에 걸쳐 통합된 메타데이터 리포지토리를 제공합니다. AWS Glue 크롤러는 Amazon S3의 소스 데이터에서 스키마를 자동으로 추론하고 관련 메타데이터를 데이터 카탈로그에 저장할 수 있습니

다. 데이터 카탈로그에 대한 자세한 내용은 AWS Glue 개발자 안내서의 Glue 데이터 카탈로그 채우기를 참조하세요. AWS

AWS Glue에는 별도의 요금이 적용됩니다. 데이터 카탈로그의 메타데이터 저장 및 액세스에 대한 월별 요금, Glue ETL 작업 및 AWS 크롤러 런타임에 대해 분당 청구되는 시간당 요금, 프로비저닝된 각 개발 엔드포인트에 대해 분당 청구되는 시간당 요금이 있습니다. 데이터 카탈로그에는 최대 100만 개의 객체를 무료로 저장할 수 있습니다. 100만 개 이상의 객체를 저장하는 경우, 100만 이상의 객체 100,000개마다 1USD가 부과됩니다. 데이터 카탈로그의 객체로는 테이블, 파티션 또는 데이터베이스가 있습니다. 자세한 내용은 Glue 요금을 참조하십시오.

> **Important**
>
> 2017년 8월 14일 이전에 Amazon Athena 또는 Amazon Redshift Spectrum을 사용하여 테이블을 생성한 경우 데이터베이스 및 테이블은 AWS Glue 데이터 카탈로그와 별도의 Athena 관리형 카탈로그에 저장됩니다. Amazon EMR을 이러한 테이블과 통합하려면 AWS Glue 데이터 카탈로그로 업그레이드해야 합니다. 자세한 내용은 Amazon Athena 사용 설명서의 AWS Glue 데이터 카탈로그로 업그레이드를 참조하세요.

Glue 데이터 카탈로그를 AWS 메타스토어로 지정

AWS Management Console AWS CLI또는 Amazon EMR API를 사용하여 AWS Glue 데이터 카탈로그를 메타스토어로 지정할 수 있습니다. CLI 또는 API를 사용할 때 Presto에 대한 구성 분류를 사용하여 데이터 카탈로그를 지정합니다. 또한 Amazon EMR 5.16.0 이상에서는 구성 분류를 사용하여 다른에서 데이터 카탈로그를 지정할 수 있습니다 AWS 계정. 콘솔을 사용할 때 고급 옵션 또는 빠른 옵션을 사용하여 데이터 카탈로그를 지정할 수 있습니다.

Console

새 콘솔을 사용하여 AWS Glue 데이터 카탈로그를 Hive 메타스토어로 지정하려면

1. 에 로그인 AWS Management Console하고 https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 왼쪽 탐색 창의 EMR on EC2에서 클러스터를 선택하고 클러스터 생성을 선택합니다.
3. 애플리케이션 번들에서 Presto를 선택합니다.
4. AWS Glue 데이터 카탈로그 설정에서 Presto 테이블 메타데이터에 사용 확인란을 선택합니다.
5. 클러스터에 적용할 다른 옵션을 선택합니다.
6. 클러스터를 시작하려면 클러스터 생성을 선택합니다.

CLI

를 사용하여 AWS Glue 데이터 카탈로그를 기본 Hive 메타스토어로 지정하려면 AWS CLI

클러스터를 생성할 때 다음 구성 분류를 지정하는 방법에 대한 예제는 애플리케이션 구성 섹션을 참조하세요.

Amazon EMR 5.16.0 이상

- 아래 JSON 예제에서와 같이 `hive.metastore` 속성을 `glue`으로 설정합니다.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

다른에서 데이터 카탈로그를 지정하려면 다음 JSON 예제와 같이 `hive.metastore.glue.catalogid` 속성을 AWS 계정추가합니다. *`acct-id`*를 데이터 카탈로그의 AWS 계정으로 바꿉니다. 다른 AWS 계정 에서 데이터 카탈로그를 사용하면 Amazon EMR 버전 5.15.0 및 이전 버전을 사용할 수 없습니다.

```
[
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore": "glue",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
]
```

Amazon EMR 5.10.0부터 5.15.0까지

아래 JSON 예제에서와 같이 `hive.metastore.glue.datacatalog.enabled` 속성을 `true`로 설정합니다.

```
[
```

```
  {
    "Classification": "presto-connector-hive",
    "Properties": {
      "hive.metastore.glue.datacatalog.enabled": "true"
    }
  }
]
```

PrestoSQL(Trino)을 사용하는 Amazon EMR 6.1.0 이상

EMR 버전 6.1.0부터 PrestoSQL도 Glue를 기본 Hive 메타스토어로 지원합니다. 다음 JSON 예제와 같이 `prestosql-connector-hive` 구성 분류를 사용하여 `hive.metastore` 속성을 `glue`로 설정합니다.

Amazon EMR 버전 6.4.0 이상에서는 PrestoSQL 대신 Trino라는 새 이름을 사용합니다. Trino를 사용하는 경우 다음 구성 분류에서 *`prestosql-connector-hive`*를 `trino-connector-hive`로 바꿉니다.

```
[
  {
    "Classification": "prestosql-connector-hive",
    "Properties": {
      "hive.metastore": "glue"
    }
  }
]
```

장기 실행 중인 클러스터에서 메타스토어를 전환하려면, 마스터 노드를 연결하고 `/etc/presto/conf/catalog/hive.properties` 파일의 속성 값을 직접 편집하고 Presto 서버(`sudo restart presto-server`)를 재시작하여, 해당 값을 릴리스 버전에 적합하도록 수동 설정할 수 있습니다. Amazon EMR 5.15.0 이하 버전에서 이 메서드를 사용하는 경우, `hive.table-statistics-enabled`가 `false`로 설정되어 있는지 확인합니다. 5.16.0 이상 릴리스 버전을 사용하는 경우 이러한 설정이 필요하지 않습니다. 그러나 나중에 테이블 및 파티션 통계는 지원되지 않습니다.

IAM 권한

클러스터의 EC2 인스턴스 프로파일에는 AWS Glue 작업에 대한 IAM 권한이 있어야 합니다. 또한 AWS Glue Data Catalog 객체에 대한 암호화를 활성화하는 경우 역할은 암호화에 AWS KMS key 사용되는를 암호화, 복호화 및 생성할 수 있어야 합니다.

AWS Glue 작업에 대한 권한

Amazon EMR에 대한 기본 EC2 인스턴스 프로파일을 사용할 경우 어떤 작업도 필요하지 않습니다. 에 연결된 `AmazonElasticMapReduceforEC2Role` 관리형 정책은 필요한 모든 AWS Glue 작업을 `EMR_EC2_DefaultRole` 허용합니다. 그러나 사용자 지정 EC2 인스턴스 프로파일 및 권한을 지정하는 경우 적절한 AWS Glue 작업을 구성해야 합니다. `AmazonElasticMapReduceforEC2Role` 관리형 정책을 시작점으로 사용합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 [클러스터 EC2 인스턴스의 서비스 역할(EC2 인스턴스 프로파일)](#)을 참조하세요.

AWS Glue 데이터 카탈로그 암호화 및 복호화 권한

인스턴스 프로파일에 키를 사용하여 데이터를 암호화하고 복호화할 수 있는 권한이 필요합니다. 다음 두 설명이 모두 적용되는 경우에는 이러한 권한을 구성할 필요가 없습니다.

- AWS Glue용 관리형 키를 사용하여 AWS Glue Data Catalog 객체에 대한 암호화를 활성화합니다.
- Glue 데이터 카탈로그 AWS 계정 와 동일한 AWS 에 있는 클러스터를 사용합니다.

그렇지 않으면 EC2 인스턴스 프로파일에 연결된 권한 정책에 다음 명령을 추가해야 합니다.

```
[
    {
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
```

```
]
```

Glue 데이터 카탈로그 암호화에 대한 자세한 내용은 AWS AWS Glue 개발자 안내서의 데이터 카탈로그 암호화를 참조하세요.

리소스 기반 권한

Amazon EMR에서 AWS Glue를 Hive, Spark 또는 Presto와 함께 사용하는 경우 AWS Glue는 리소스 기반 정책을 지원하여 데이터 카탈로그 리소스에 대한 액세스를 제어합니다. 이러한 리소스에는 데이터베이스, 테이블, 연결 및 사용자 정의 기능이 포함됩니다. 자세한 내용은 AWS Glue 개발자 안내서에서 AWS Glue 리소스 정책을 참조하세요.

리소스 기반 정책을 사용하여 Amazon EMR 내에서 AWS Glue에 대한 액세스를 제한하는 경우 권한 정책에서 지정하는 보안 주체는 클러스터 생성 시 지정된 EC2 인스턴스 프로파일과 연결된 역할 ARN이어야 합니다. 예를 들어, 카탈로그에 연결된 리소스 기반 정책의 경우 다음 예에 표시된 형식을 사용하여 클러스터 EC2 인스턴스에 대한 기본 서비스 역할의 ARN(*EMR_EC2_DefaultRole*)을 `Principal`로 지정할 수 있습니다.

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*는 AWS Glue 계정 ID와 다를 수 있습니다. 그러면 다른 계정의 EMR 클러스터에서 액세스할 수 있습니다. 각각 다른 계정에서 여러 보안 주체를 지정할 수 있습니다.

AWS Glue 데이터 카탈로그 사용 시 고려 사항

Presto에서 AWS Glue 데이터 카탈로그를 메타스토어로 사용할 때는 다음 항목을 고려하세요.

- AWS Glue 내에서 테이블 이름을 바꾸는 것은 지원되지 않습니다.
- `LOCATION`을 지정하지 않고 Hive 테이블을 생성하면 테이블 데이터가 `hive.metastore.warehouse.dir` 속성을 통해 지정된 위치에 저장됩니다. 기본적으로 HDFS에 있는 위치입니다. 다른 클러스터가 테이블에 액세스해야 하는 경우, 테이블을 생성한 클러스터에 대한 적절한 권한이 없으면 실패합니다. 또한 HDFS 스토리지는 일시적이기 때문에, 클러스터가 종료되는 경우 테이블 데이터가 손실되고 테이블을 다시 생성해야 합니다. Glue를 사용하여 Hive 테이블을 생성할 때 Amazon S3`LOCATION`에서 AWS 를 지정하는 것이 좋습니다. 또는 `hive-site` 구성 분류를 사용하여 모든 Hive 테이블에 적용되는 Amazon S3에서 `hive.metastore.warehouse.dir`에 대한 위치를 지정할 수도 있습니다. HDFS 위치에 테이블이 생성되고 테이블을 생성한 클러스터가 여전히 실행 중인 경우 AWS Glue 내에서 테이블 위치를 Amazon S3로 업데이트할 수 있습니다. 자세한 내용은 Glue 개발자 안내서의 AWS Glue 콘솔에서 테이블 작업을 참조하세요. AWS

- 따옴표와 아포스트로피가 포함된 파티션 값은 지원되지 않습니다(예: `PARTITION (owner="Doe's").`).
- 열 통계는 emr-5.31.0 이상에서 지원됩니다.
- Hive 권한 부여 사용은 지원되지 않습니다. 대안으로 AWS Glue 리소스 기반 정책 사용을 고려합니다. 자세한 내용은 AWS Glue 데이터 카탈로그에 대한 Amazon EMR 액세스를 위한 리소스 기반 정책 사용을 참조하세요.

S3 Select Pushdown과 함께 Presto를 사용하여 성능 향상

> **Important**
>
> 신규 고객은 더 이상 Amazon S3 Select를 사용할 수 없습니다. Amazon S3 Select 기존 고객은 평소처럼 이 기능을 계속 사용할 수 있습니다. 자세히 알아보기

Amazon EMR 릴리스 버전 5.18.0 이상에서는 Amazon EMR 기반 Presto에서 S3 Select 푸시다운을 사용할 수 있습니다. 이 기능을 사용하면 프로젝션 작업의 계산 작업(예: `SELECT`)과 조건자 작업(예: `WHERE`)을 Amazon S3로 '푸시다운'할 수 있습니다. 이를 통해 쿼리는 Amazon S3에서 필요한 데이터만 검색할 수 있으므로 일부 애플리케이션에서 성능을 향상시키고 Amazon EMR과 Amazon S3 사이에 전송되는 데이터의 양을 줄일 수 있습니다.

내 애플리케이션에 S3 Select Pushdown이 적합한가요?

S3 Select Pushdown 사용 여부에 관계없이 애플리케이션을 벤치마킹하여 애플리케이션에 적합한지 여부를 확인하는 것이 좋습니다.

다음 지침을 사용하여 애플리케이션과 함께 S3 Select를 사용할 수 있는지 확인하십시오.

- 쿼리가 원본 데이터 세트 중 반 이상을 필터링합니다.
- 쿼리 필터 조건자는 Presto 및 S3 Select에서 지원하는 데이터 형식의 열을 사용합니다. 타임스탬프, real 및 double 데이터 형식은 S3 Select Pushdown에서 지원되지 않습니다. 숫자 데이터에는 10진수 데이터 형식을 사용하는 것이 좋습니다. S3 Select에 지원되는 데이터 유형에 대한 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 데이터 유형을 참조하세요.
- Amazon S3 및 Amazon EMR 클러스터 간 네트워크 연결의 전송 속도와 가용 대역폭이 양호합니다. Amazon S3는 HTTP 응답을 압축하지 않으므로 압축된 입력 파일의 경우 응답 크기가 증가할 수 있습니다.

고려 사항 및 제한 사항

- CSV 형식으로 저장된 개체만 지원됩니다. 객체는 압축 해제되거나 gzip 또는 bzip2로 선택적으로 압축될 수 있습니다.
- `AllowQuotedRecordDelimiters` 속성이 지원되지 않습니다. 이 속성이 지정되면 쿼리가 실패합니다.
- 고객 제공 암호화 키를 사용하는 Amazon S3 서버 측 암호화(SSE-C) 및 클라이언트 측 암호화는 지원되지 않습니다.
- S3 Select Pushdow은 ORC 또는 Parquet과 같은 열 또는 압축 파일 형식 대신 사용할 수 없습니다.
- Amazon S3 Select는 Presto 및 Trino에 대한 리전 간 쿼리를 지원하지 않습니다.

Presto 또는 Trino를 사용하여 S3 Select Pushdown 활성화

Amazon EMR에서 Presto에 대해 S3 Select Pushdown을 활성화하려면 아래 예제와 같이 `hive.s3select-pushdown.enabled` 구성 분류를 사용하여 `presto-connector-hive`를 `true`로 설정합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오. hive.s3select-pushdown.max-connections 값도 설정해야 합니다. 대부분의 애플리케이션에서는 기본 설정 *500*이 적합합니다. 자세한 내용은 아래 hive.s3select-pushdown.max-connections 이해 및 튜닝 섹션을 참조하세요.

EMR 버전 6.1.0~6.3.0 기반 PrestoSQL의 경우 아래 예제에서 *presto-connector-hive*를 `prestosql-connector-hive`로 바꿉니다.

Amazon EMR 버전 6.4.0 이상에서는 PrestoSQL 대신 Trino라는 새 이름을 사용합니다. Trino를 사용하는 경우 아래 예제에서 *presto-connector-hive*를 `trino-connector-hive`로 바꿉니다.

```
[
    {
        "classification": "presto-connector-hive",
        "properties": {
            "hive.s3select-pushdown.enabled": "true",
            "hive.s3select-pushdown.max-connections": "500"
        }
    }
]
```

hive.s3select-pushdown.max-connections 이해 및 튜닝

기본적으로 Presto는 파일 시스템으로 EMRFS를 사용합니다. `emrfs-site` 구성 분류의 `fs.s3.maxConnections` 설정은 Presto용 EMRFS를 통해 Amazon S3에 허용되는 최대 클라이언트 연결을 지정합니다. 기본적으로 이 값은 500입니다. S3 Select Pushdown은 조건자 작업에 대해 Amazon S3에 액세스할 때 EMRFS를 우회합니다. 이 경우 `hive.s3select-pushdown.max-connections` 값은 작업자 노드에서 이러한 작업에 허용되는 최대 클라이언트 연결 수를 결정합니다. 하지만 Presto가 시작하고 푸시다운되지 않은 Amazon S3에 대한 모든 요청(예: GET 작업)은 계속해서 `fs.s3.maxConnections`의 값으로 관리됩니다.

애플리케이션에서 "풀에서의 연결 대기 시간이 초과되었습니다"라는 오류가 발생하면 `hive.s3select-pushdown.max-connections` 및 `fs.s3.maxConnections` 값을 모두 늘립니다.

데이터베이스 커넥터 추가

클러스터를 생성할 때 구성 분류를 사용하여 JDBC 커넥터 속성을 구성할 수 있습니다. 구성 분류는 `presto-connector-postgresql`과 같이 `presto-connector`로 시작합니다. 사용 가능한 구성 분류는 Amazon EMR 릴리스 버전에 따라 다릅니다. 최신 릴리스 버전에서 사용할 수 있는 구성 분류에 대해서는 Amazon EMR 5.36.2에 대한 [the section called “5.36.2 구성 분류”](#) 섹션을 참조하세요. 다른 버전의 Amazon EMR을 사용하는 경우 구성 분류에 대해서는 [Amazon EMR 5.x 릴리스 버전](#) 섹션을 참조하세요. 각 커넥터로 구성할 수 있는 속성에 대한 자세한 내용은 https://prestodb.io/docs/current/connector.html을 참조하십시오.

Example - PostgreSQL JDBC 커넥터를 사용하여 클러스터 구성

PostgreSQL 커넥터가 설치 및 구성된 클러스터를 시작하려면 먼저 다음 콘텐츠를 포함하는 구성 분류(예:`myConfig.json`)를 지정하는 JSON 파일을 생성하고 로컬에 저장합니다.

Presto 설명서에서 [PostgreSQL connector](#) 주제에 표시된 대로 연결 속성을 설정에 맞게 바꿉니다.

```
[
  {
    "Classification": "presto-connector-postgresql",
    "Properties": {
      "connection-url": "jdbc:postgresql://example.net:5432/database",
      "connection-user": "MYUSER",
      "connection-password": "MYPASS"
    },
    "Configurations": []
```

```
  }
]
```

클러스터를 생성할 때 다음 예와 같이 `--configurations` 옵션을 사용하여 JSON 파일의 경로를 참조하십시오. 여기서 `myConfig.json`은 명령을 실행하는 디렉터리와 같습니다.

```
aws emr create-cluster --name PrestoConnector --release-label emr-5.36.2 --instance-type m5.xlarge \
--instance-count 2 --applications Name=Hadoop Name=Hive Name=Pig Name=Presto \
--use-default-roles --ec2-attributes KeyName=myKey \
--log-uri s3://amzn-s3-demo-bucket/logs --enable-debugging \
--configurations file://myConfig.json
```

Amazon EMR 기반 Presto를 통해 SSL 및 TLS 사용과 LDAPS 구성

Amazon EMR 릴리스 버전 5.6.0 이상으로 SSL 및 TLS를 활성화하면 Presto 노드 사이에서 안전한 내부 통신이 가능합니다. 이렇게 하려면 전송 중 데이터 암호화를 위한 보안 구성을 설정합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 암호화 옵션 및 보안 구성을 사용하여 클러스터 보안 설정을 참조하세요.

전송 중 데이터 암호화를 통해 보안 구성을 사용하면 Amazon EMR이 Presto에 대해 다음을 수행합니다.

- 전송 중 데이터 암호화를 위해 지정하는 암호화 아티팩트 또는 아티팩트를 Pedro 클러스터 전체로 분산시킵니다. 자세한 내용은 전송 중 데이터 암호화에 대한 인증서 제공을 참조하세요.
- Presto에 대한 `presto-config`에 상응하는 `config.properties` 구성 분류를 사용하여 다음 속성을 설정합니다.
  - 모든 노드에서 `http-server.http.enabled`를 `false`로 설정하여 HTTPS를 사용할 수 있도록 HTTP를 비활성화합니다. 이를 위해 전송 중 암호화를 위한 보안 구성을 설정할 때 퍼블릭 및 프라이빗 DNS에 사용할 수 있는 인증서를 제공해야 합니다. 이를 위한 한 가지 방법은 여러 도메인을 지원하는 주체 대체 이름(SAN) 인증서를 사용하는 것입니다.
  - `http-server.https.*` 값을 설정합니다. 구성에 대한 자세한 내용은 Presto 설명서에서 LDAP authentication을 참조하세요.

또한 Amazon EMR 릴리스 버전 5.10.0 이상을 사용하여, HTTPS를 사용하는 Presto 코디네이터에 대한 클라이언트 연결을 위해 LDAP 인증을 설정할 수 있습니다. 이 설정은 안전한 LDAP(LDAPS)를 사용합니다. LDAP 서버에서 TLS를 활성화해야 하며 Presto 클러스터는 전송 중 데이터 암호화가 활성

화된 상태에서 보안 구성을 사용해야 합니다. 추가 구성이 필요합니다. 사용하는 Amazon EMR 릴리스 버전에 따라 구성 옵션이 달라집니다. 자세한 내용은 Amazon EMR 기반 Presto에 대한 LDAP 인증 사용 단원을 참조하십시오.

Amazon EMR 기반 Presto는 기본적으로 내부 HTTPS에 포트 8446을 사용합니다. 내부 통신에 사용되는 포트는 Presto 조정자에 대한 클라이언트 HTTPS 액세스에 사용되는 포트와 동일해야 합니다. `http-server.https.port` 구성 분류의 `presto-config` 속성이 포트를 지정합니다.

Amazon EMR 기반 Presto에 대한 LDAP 인증 사용

이번 섹션의 단계를 따르면 LDAP를 구성할 수 있습니다. 예제의 단계 및 추가 정보 링크를 참조하십시오.

LADP 인증 구성 단계

- 1단계: LADP 서버에 대한 정보를 수집하고 서버 인증서를 Amazon S3로 복사
- 2단계: 보안 구성 설정
- 3단계: LDAP에 대한 Presto 속성을 통한 구성 JSON 생성
- 4단계: 스크립트를 생성하여 LADP 서버 인증서를 복사하여 Amazon S3로 업로드
- 5단계: 클러스터 생성

1단계: LADP 서버에 대한 정보를 수집하고 서버 인증서를 Amazon S3로 복사

LDAP 서버로부터 다음 섹션의 정보 및 항목을 수집하여 LDAP 인증을 구성해야 합니다.

LDAP 서버의 IP 주소 또는 호스트 이름

Amazon EMR 프라이머리 노드의 Presto 코디네이터가 지정된 IP 주소 또는 호스트 이름의 LDAP 서버에 도달할 수 있어야 합니다. 기본적으로 Presto는 LDAPS를 사용하여 포트 636을 통해 LDAP 서버와 통신합니다. LDAP 실행에 사용자 지정 포트가 필요한 경우 Amazon EMR 5.16.0 이상의 `ldap.url` 속성을 사용하거나 이전 버전의 `authentication.ldap.url`을 사용하여 포트를 지정할 수 있습니다. 636의 `presto-config` 구성 분류 예에 표시된 대로 사용자 지정 포트3단계: LDAP에 대한 Presto 속성을 통한 구성 JSON 생성을 대체합니다. 모든 방화벽 및 보안 그룹에서 포트 636 (또는 사용자 지정 포트)의 인바운드 및 아웃바운드 트래픽을 허용하고 또한 내부 클러스터 통신에 사용되는 포트 8446 (또는 사용자 지정 포트)에 대해서도 허용하는지 확인합니다.

LDAP 서버 인증서

인증서 파일을 Amazon S3의 안전한 위치에 업로드해야 합니다. 자세한 내용은 Amazon Simple Storage Service 사용 설명서에서 S3 버킷에 파일 및 폴더를 업로드하려면 어떻게 해야 하나요? 섹션

을 참조하세요. 클러스터를 시작할 때 이 인증서를 Amazon S3에서 클러스터의 각 노드로 복사하는 부트스트랩 작업을 생성합니다. In [4단계: 스크립트를 생성하여 LADP 서버 인증서를 복사하여 Amazon S3로 업로드](). 예제 인증서는 *s3://amzn-s3-demo-bucket/ldap_server.crt*입니다.

익명 바인딩을 위한 LDAP 서버 설정

익명 바인딩을 PrestoDB에서 비활성화한 경우 LDAP 서버와의 바인딩이 허가된 계정의 사용자 ID(UID) 및 암호가 있어야 PrestoDB 서버가 연결을 설정할 수 있습니다. `internal-communication.authentication.ldap.user` 구성 분류에서 `internal-communication.authentication.ldap.password` 및 `presto-config` 속성을 사용하여 UID 및 암호를 지정해야 합니다. Amazon EMR 5.10.0은 이러한 설정을 지원하지 않으므로, 이 릴리스 버전을 사용하는 경우 LDAP 서버에서 익명 바인딩이 지원되어야 합니다.

참고로 Trino에는 익명 바인딩 구성이 필요하지 않습니다.

LDAP 서버에서 익명 바인딩의 상태를 가져오려면

- 다음 예와 같이 Linux 클라이언트에서 [ldapwhoami]() 명령을 사용합니다.

```
ldapwhoami -x -H ldaps://LDAPServerHostNameOrIPAddress
```

  익명 바인딩이 허용되지 않는 경우 명령이 다음을 반환합니다.

```
ldap_bind: Inappropriate authentication (48)
additional info: anonymous bind disallowed
```

특정 계정에 단순 인증을 사용하는 LDAP 서버에 대한 액세스 권한이 있는지 확인하려면

- 다음 예와 같이 Linux 클라이언트에서 [ldapwhoami]() 명령을 사용합니다. 이 예제에서는 가상의 호스트 이름 *ip-xxx-xxx-xxx-xxx.ec2.internal*의 EC2 인스턴스에서 실행 중인 Open LDAP 서버에 저장된 가상의 사용자, *presto*를 사용합니다. 사용자는 조직 단위(OU) *admins* 및 암호 *123456*와 연결되어 있습니다.

```
ldapwhoami -x -w "123456" -D uid=presto,ou=admins,dc=ec2,dc=internal -H ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal
```

  유효한 계정이며 적절한 권한이 부여되어 있다면 명령이 다음을 반환합니다.

```
dn:uid=presto,ou=admins,dc=ec2,dc=internal
```

3단계: LDAP에 대한 Presto 속성을 통한 구성 JSON 생성의 구성 예에는 명확한 이해를 위해 이 계정이 포함되나, 이 계정을 지원하지 않는 5.10.0 예는 예외입니다. LDAP 서버가 익명 바인딩을 사용하는 경우 `internal-communication.authentication.ldap.user`와 `internal-communication.authentication.ldap.password` 이름/값 페어를 제거합니다.

Presto 사용자용 LDAP 고유 이름(DN)

Presto에 대해 LDAP 구성을 지정하면 `${USER}`로 이루어진 바인드 패턴이 조직 단위(OU) 및 추가적인 도메인 구성 요소(DC)와 함께 지정됩니다. Presto는 암호 인증 과정에서 `${USER}`를 실제 사용자 ID(UID)로 대체하여 이 바인드 패턴을 통해 지정한 고유 이름(DN)과 대조합니다. 적격 사용자가 소속된 OU(조직 단위) 및 DC(도메인 구성 요소)가 필요합니다. 예를 들어 `admins` 도메인의 `corp.example.com` OU에 소속된 사용자에게 권한을 부여하여 Presto에 인증하려면 `${USER},ou=admins,dc=corp,dc=example,dc=com`를 사용자 바인드 패턴으로 지정해야 합니다.

> **Note**
>
> 를 사용할 때를 실제 사용자 ID(UID)`${USER}`로 바꾸려면 Fn::Sub 함수를 사용해야 AWS CloudFormation합니다. 자세한 내용은 AWS CloudFormation 사용 설명서에서 Fn::Sub 주제를 참조하세요.

Amazon EMR 5.10.0을 사용하는 경우 해당 패턴을 1개만 지정할 수 있습니다. Amazon EMR 5.11.0 이상을 사용하는 경우 여러 개의 패턴을 콜론(:)으로 구분하여 지정할 수 있습니다. Presto에 인정받으려는 사용자를 첫 번째 패턴과 대조한 다음 두 번째 패턴과 대조한 후에 또 그와 같은 과정을 반복하는 식으로 진행합니다. 예제는 3단계: LDAP에 대한 Presto 속성을 통한 구성 JSON 생성 섹션을 참조하세요.

2단계: 보안 구성 설정

전송 중 데이터 암호화가 활성화된 보안 구성을 생성합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 보안 구성 생성을 참조하세요. 전송 중 데이터 암호화를 설정할 때 제공하는 암호화 아티팩트는 Presto 노드 간의 내부 통신을 암호화하는 데 사용합니다. 자세한 내용은 전송 중 데이터 암호화에 대한 인증서 제공을 참조하세요. LDAP 서버 인증서는 Presto 클라우드에 대한 클라이언트 연결을 인증하는 데 사용합니다.

3단계: LDAP에 대한 Presto 속성을 통한 구성 JSON 생성

presto-config 구성 분류를 사용하여 LDAP에 대한 Presto 속성을 설정합니다. presto-config의 형식과 콘텐츠는 Amazon EMR 릴리스 버전과 Presto 설치(PrestoDB 또는 Trino)에 따라 조금씩 다릅니다. 이 섹션의 뒷부분에서 구성 차이의 예를 제공합니다. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

다음 단계에서는 사용자가 JSON 데이터를 *MyPrestoConfig.json* 파일에 저장했다고 가정합니다. 콘솔을 사용하는 경우 클러스터를 생성할 때 파일을 참조할 수 있도록 파일을 Amazon S3의 안전한 위치에 업로드해야 합니다. 를 사용하는 경우 파일을 로컬에서 참조 AWS CLI할 수 있습니다.

Example PrestoSQL(Trino)을 지원하는 Amazon EMR 6.1.0 이상

다음 예제에서는 1단계: LADP 서버에 대한 정보를 수집하고 서버 인증서를 Amazon S3로 복사의 LDAP 호스트 이름을 사용하여 바인딩을 위해 LDAP 서버에 인증합니다. 2가지 사용자 바인드 패턴을 지정하여 LDAP 서버의 admins OU 및 datascientists OU 내에 있는 사용자가 사용자로서 Trino 서버에 인증할 수 있는 자격이 있음을 보여줍니다. 바인드 패턴은 콜론으로 구분됩니다(:).

Amazon EMR 버전 6.4.0 이상에서는 PrestoSQL 대신 Trino라는 새 이름을 사용합니다. Trino를 사용하는 경우 다음 구성 분류에서 *prestosql-config*를 trino-config로, *prestosql-password-authenticator*를 trino-password-authenticator로 바꿉니다.

```
[
   {
      "Classification":"prestosql-config",
      "Properties":{
         "http-server.authentication.type":"PASSWORD"
      }
   },
   {
      "Classification":"prestosql-password-authenticator",
      "Properties":{
         "password-authenticator.name":"ldap",
         "ldap.url":"ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
         "ldap.user-bind-pattern": "uid=${USER},ou=admins,dc=ec2,dc=internal:uid=
${USER},ou=datascientists,dc=ec2,dc=internal"
      }
   }
]
```

Example Amazon EMR 5.16.0 이상

다음 예에서는 LDAP 사용자 ID 및 암호와 [1단계: LADP 서버에 대한 정보를 수집하고 서버 인증서를 Amazon S3로 복사](#)의 LDAP 호스트 이름을 사용하여 LDAP 서버에 바인딩을 인증합니다. 2가지 사용자 바인드 패턴을 지정하여 LDAP 서버의 `admins` OU 및 `datascientists` OU 내에 있는 사용자가 사용자로서 Presto 서버에 인증할 수 있는 자격이 있음을 보여줍니다. 바인드 패턴은 콜론으로 구분됩니다(`:`).

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "PASSWORD"
                }
        },
        {
                "Classification": "presto-password-authenticator",
                "Properties": {
                        "password-authenticator.name": "ldap",
                        "ldap.url": "ldaps://ip-xxx-xxx-xxx-xxx.ec2.internal:636",
                        "ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
                }
        }]
```

Example Amazon EMR 5.11.0부터 5.15.0까지

이러한 릴리스 버전에서는 `presto-config` 구성 분류의 형식이 조금씩 다릅니다. 다음 예는 이전의 예와 동일한 파라미터를 지정합니다.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
                        "authentication.ldap.user-bind-pattern": "uid=
${USER},ou=admins,dc=ec2,dc=internal:uid=${USER},ou=datascientists,dc=ec2,dc=internal",
                        "internal-communication.authentication.ldap.user": "presto",
                        "internal-communication.authentication.ldap.password": "123456"
                }
```

```
        }]
```

Example Amazon EMR 5.10.0

Amazon EMR 5.10.0은 익명 바인딩만을 지원하므로 그러한 항목은 생략됩니다. 또한 단일 바인드 패턴만 지정할 수 있습니다.

```
[{
        "Classification": "presto-config",
                "Properties": {
                        "http-server.authentication.type": "LDAP",
                        "authentication.ldap.url": "ldaps://ip-xxx-xxx-xxx-
xxx.ec2.internal:636",
                        "ldap.user-bind-pattern": "uid=
${USER},ou=prestousers,dc=ec2,dc=internal"
                }
        }]
```

4단계: 스크립트를 생성하여 LADP 서버 인증서를 복사하여 Amazon S3로 업로드

인증서 파일을 클러스터의 각 노드에 복사하고 키 스토어에 추가하는 스크립트를 생성합니다. 텍스트 편집기를 사용하여 스크립트를 생성하고 저장한 다음 Amazon S3에 업로드합니다. 5단계: 클러스터 생성에서 스크립트 파일은 *s3://amzn-s3-demo-bucket/LoadLDAPCert.sh*로 참조됩니다.

다음 예제 스크립트는 기본 키 스토어 암호로 *changeit*을 사용합니다. 클러스터를 생성한 후에 마스터 노드에 연결하고 keytool 명령을 사용하여 키 스토어 암호를 변경하는 것이 좋습니다.

```
#!/bin/bash
aws s3 cp s3://amzn-s3-demo-bucket/ldap_server.crt .
sudo keytool -import -keystore /usr/lib/jvm/jre-1.8.0-openjdk.x86_64/lib/security/
cacerts -trustcacerts -alias ldap_server -file ./ldap_server.crt -storepass changeit -
noprompt
```

5단계: 클러스터 생성

클러스터를 생성할 때 Amazon EMR을 설치할 Presto와 기타 애플리케이션을 지정합니다. 또한 다음 예에서는 JSON 내의 구성 분류 속성을 참조하지만, 구성 분류 인라인을 지정할 수도 있습니다.

Amazon EMR 관리 콘솔을 사용하여 LDAP 인증을 통해 Presto 클러스터를 생성하는 방법

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. 설치할 Amazon EMR용 다른 애플리케이션과 함께 Presto를 선택하고 소프트웨어 구성에서 사용할 Amazon EMR 릴리스를 선택합니다. LDAP 인증은 Amazon EMR 5.10.0 이상에서만 지원됩니다.
4. 소프트웨어 설정 편집에서 S3로부터 JSON 로드를 선택하고 3단계: LDAP에 대한 Presto 속성을 통한 구성 JSON 생성에서 생성한 JSON 구성 파일의 Amazon S3 위치를 입력한 후 다음을 선택합니다.
5. 클러스터 하드웨어 및 네트워킹을 구성한 다음 다음을 선택합니다.
6. 부트스트랩 작업을 선택합니다. 부트스트랩 작업 추가에서 사용자 지정 작업을 선택한 다음 구성 및 추가를 선택합니다.
7. 부트스트랩 작업의 이름을 입력하고 4단계: 스크립트를 생성하여 LADP 서버 인증서를 복사하여 Amazon S3로 업로드에서 생성한 스크립트 위치(예: s3://amzn-s3-demo-bucket/LoadLDAPCert.sh)를 입력한 다음, 추가를 선택합니다.
8. 일반 옵션의 태그 및 추가 옵션에서 애플리케이션에 적합한 설정을 선택한 후 다음을 선택합니다.
9. 인증 및 암호화를 선택하고 2단계: 보안 구성 설정에서 생성한 보안 구성을 선택합니다.
10. 애플리케이션에 적합한 다른 보안 옵션을 선택한 다음 클러스터 생성을 선택합니다.

AWS CLI를 사용하여 LDAP 인증을 통해 Presto 클러스터를 생성하려면

- `aws emr create-cluster` 명령을 사용합니다. 최소한 Presto 애플리케이션과 Presto 구성 분류, 부트스트랩 스크립트 및 이전 단계에서 생성한 보안 구성을 지정해야 합니다. 다음 예에서는 명령 실행에 사용된 것과 동일한 디렉터리에 저장된 JSON 파일인 구성 파일을 참조합니다. 한편 부트스트랩 스크립트는 Amazon S3에 저장되어야 합니다. 다음 예에는 `s3://amzn-s3-demo-bucket/LoadLDAPCert.sh`가 사용됩니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster  --applications Name=presto --release-label emr-5.16.0 \
--use-default-roles --ec2-attributes KeyName=MyKeyPair,SubnetId=subnet-1234ab5
 \ --instance-count 3 --instance-type m5.xlarge --region us-west-2 --name
 "MyPrestoWithLDAPAuth" \
--bootstrap-actions Name="Distribute LDAP server cert",Path="s3://amzn-s3-demo-
bucket/LoadLDAPCert.sh" \
--security-configuration MyPrestoLDAPSecCfg --configurations file://
MyPrestoConfig.json
```

Presto 엄격한 모드 활성화

경우에 따라 장기 실행 쿼리로 인해 비용이 증가하고 Amazon EMR이 더 많은 클러스터 리소스를 사용할 수 있습니다. 이렇게 하면 클러스터의 다른 워크로드에서 리소스를 가져옵니다. Amazon EMR 버전 6.8 이상에서는 다음과 같은 유형의 장기 실행 쿼리를 거부하거나 경고하는 엄격한 모드 기능을 사용할 수 있습니다.

- 파티셔닝된 열에 조건자가 없는 쿼리는 테이블에서 대량의 데이터를 스캔하게 됩니다.
- 두 개의 대형 테이블 간 교차 조인이 있는 쿼리
- 많은 행을 제한 없이 정렬하는 쿼리

Presto가 쿼리 계획을 완전히 최적화한 후 엄격한 모드가 실행됩니다. 엄격한 모드를 사용하고 쿼리 요구 사항에 맞게 사용자 지정하려면 다음과 같은 방법으로 Presto를 구성할 수 있습니다.

엄격한 모드를 위한 Presto 구성

설정	설명	Default
`strict-mode-enabled`	엄격한 모드를 켜고 끕니다. 값이 `true`이면 엄격한 모드가 켜져 있음을 나타냅니다.	`false`
`strict-mode-fail-query`	엄격한 모드에서 장기 실행 쿼리가 감지되면 쿼리를 거부합니다. `false`인 경우 Amazon EMR은 경고만 생성합니다.	`false`

설정	설명	Default
strict-mode-restrictions	엄격한 모드가 켜져 있을 때 적용할 제한 사항을 지정합니다. 엄격한 모드에서는 MANDATORY_PARTITION_PREDICATE, DISALLOW_CROSS_JOIN, LIMITED_SORT와 같은 제한 사항을 지원합니다.	MANDATORY_PARTITION_PREDICATE,DISALLOW_CROSS_JOIN, LIMITED_SORT

엄격한 모드를 시험해 보려면 이러한 구성을 재정의하거나 Presto 클라이언트를 사용할 때 세션 속성으로 설정할 수 있습니다.

를 사용하여 클러스터 생성 시 구성을 설정하려면 AWS Management Console

1. 클러스터 생성을 선택하고 Amazon EMR 버전 6.8.0, Presto 또는 Trino를 선택합니다. 자세한 내용은 PrestoDB 및 Trino 설치를 참조하세요.
2. 엄격한 모드의 구성 속성을 직접 지정하거나 Amazon S3에 JSON 파일을 업로드합니다. 선택적으로 메타스토어의 AWS Glue 데이터 카탈로그를 선택합니다. VPC, 서브넷, 부트스트랩 작업, 키 페어, 보안 그룹을 지정합니다. 클러스터 생성을 선택하여 클러스터를 생성합니다.
3. 클러스터의 프라이머리 노드에 로그인하고 `presto-cli` 또는 `trino-cli`를 실행합니다.
4. 쿼리를 제출합니다. 엄격한 모드는 각 쿼리를 검증하고 장기 실행 쿼리인지 확인합니다. `strict-mode-fail-query` 설정에 따라 Amazon EMR은 쿼리를 거부하거나 경고를 생성합니다.
5. 쿼리를 마치면 클러스터를 종료하고 리소스를 삭제합니다.

를 사용하여 실행 중인 클러스터에서 구성을 설정하려면 AWS CLI

1. 를 사용하여 클러스터의 프라이머리 노드에 로그인 AWS CLI 하고 `presto-cli` 또는를 실행합니다`trino-cli`.
2. 원하는 값으로 다음 명령을 실행합니다.

```
set session strict_mode_enabled = true;
set session strict_mode_fail_query = false;
set session strict_mode_restrictions = 'DISALLOW_CROSS_JOIN,LIMITED_SORT';
```

고려 사항

엄격한 모드를 사용할 때 다음 사항을 고려합니다.

- 경우에 따라 엄격한 모드에서는 리소스를 많이 소비하지 않는 단기 실행 쿼리를 거부할 수 있습니다. 예를 들어 소형 테이블에 대한 쿼리는 동적 필터링을 적용하지 않거나 내부 조인을 교차 조인으로 대체하지 않습니다. 이로 인해 쿼리에서 필수 파티션 조건자를 사용하거나 교차 조인을 허용하지 않을 수 있습니다. 이 경우 엄격한 모드에서 쿼리를 거부합니다.
- 엄격한 모드 검사는 SELECT, INSERT, CREATE TABLE AS SELECT, EXPLAIN ANALYZE 쿼리 유형에만 적용됩니다.
- Hive 커넥터에서는 필수 파티션 조건자에 대한 제한만 사용할 수 있습니다.

Presto에서 스팟 인스턴스 손실 처리

Amazon EMR에서 스팟 인스턴스를 사용하면 저렴한 비용으로 여분의 Amazon EC2 용량에서 빅 데이터 워크로드를 실행할 수 있습니다. Amazon EC2는 비용을 절감하는 대가로 2분 알림을 통해 스팟 인스턴스를 중단할 수 있습니다. 노드를 종료하면 Presto에서 오류를 반환하기까지 최대 10분이 걸릴 수 있습니다. 이로 인해 오류 보고서가 불필요하게 지연되고 재시도가 발생할 수 있습니다. 빠른 종료는 Presto가 종료된 노드를 처리하는 방식을 제어할 수 있는 기능입니다.

Presto 코디네이터의 역할은 주기적으로 상태를 폴링하여 모든 워커 노드를 추적하는 것입니다. 빠른 종료가 없으면 코디네이터는 YARN NodeManager에 각 노드의 상태를 문의하지 않습니다. 이로 인해 쿼리가 실패하기 전에 재시도 루프가 길어질 수 있습니다. 빠른 종료를 통해 Presto 코디네이터는 폴링이 호스트에 도달하지 못하는 즉시 NodeManager에서 노드 상태를 확인합니다. NodeManager에서 노드가 비활성 상태임을 표시하면 Presto는 추가 재시도를 중단하고 쿼리에 실패한 후 `NODE_DECOMMISSIONED` 오류를 반환합니다.

다음 구성 파라미터 세트를 사용하면 노드 종료 시 Presto의 동작을 제어하고 사용자 지정할 수 있습니다.

노드 장애 처리를 위한 Presto 구성

설정	설명	Default
`query.remote-task.max-backoff-duration`	코디네이터가 워커 노드에서 원격 작업 상태를 계속 가져오려고 시도하는 시간.	10분

설정	설명	Default
`query.remote-task.quick-terminate-node-failure`	코디네이터가 노드에 연결할 수 없거나 해당 노드에서 실행되는 워커에 연결할 수 없는 경우 빠른 노드 장애를 활성화합니다. `query.remote-task.terminate-on-connect-exception` 의 값에 따라 코디네이터가 노드에 도달해야 하는지 아니면 워커에 연결해야 하는지가 결정됩니다. 노드가 쿼리에 실패하고 Amazon EMR은 사용 가능한 워커 목록에서 해당 노드를 제거합니다. 이 경우 노드를 사용하여 새 쿼리를 예약할 수 없습니다. 이 값을 `false`로 설정하면 노드를 사용할 수 없는 것으로 표시하기 전에 Presto가 노드에 도달하려고 다시 시도하는 이전 동작(`query.remote-task.max-backoff-duration`)으로 돌아가고 노드에서 진행 중인 쿼리에 실패합니다.	`true`

설정	설명	Default
`query.remote-task.terminate-on-connect-exception`	호스트에 연결할 수 있지만 코디네이터가 호스트의 워커 프로세스에 연결하지 못하는 경우 Amazon EMR에서 노드를 사용해야 하는지 여부를 지정합니다. 이 값을 `true`로 설정하면 호스트에 연결할 수 없는 경우 빠른 쿼리 실패가 활성화됩니다.	`false`

단계적 서비스 해제 기능이 있는 자동 조정 사용

Amazon EMR 릴리스 버전 5.30.0 이상에는 특정 조정 작업에 대한 유예 기간을 설정하는 데 사용할 수 있는 기능이 포함되어 있습니다. 스케일 인 조정 작업 또는 자동 조정 정책 요청으로 인해 노드가 종료되기 전에 유예 기간을 통해 Presto 작업을 계속 실행할 수 있습니다. 조정 규칙에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 자동 조정 규칙 이해를 참조하세요. 단계적 서비스 해제를 포함하는 Presto 자동 조정을 사용하면 서비스 해제 중인 노드에 새 작업이 예약되지 않도록 하는 동시에 종료 제한 시간에 도달하기 전에 이미 실행 중인 작업이 완료되도록 할 수 있습니다. 실행 중인 쿼리는 노드에 대해 서비스 해제가 수행되기 전에 실행을 완료합니다. 자동 조정은 인스턴스 플릿에서 지원되지 않습니다.

자동 조정 종료 요청을 받은 후 Presto 작업을 완료하는 데 걸리는 시간을 제어할 수 있습니다. 기본적으로 Amazon EMR에 대한 종료 제한 시간은 `0`분입니다. 따라서 스케일 인 요청에 의해 필요한 경우 Amazon EMR은 노드 및 모든 Presto 작업을 즉시 종료합니다. 클러스터를 스케일 다운하기 전에 실행 중인 쿼리가 완료될 수 있도록 Amazon EMR에서 Presto 작업에 더 긴 제한 시간을 설정하려면 `presto-config` 구성 분류를 사용하여 `graceful-shutdown-timeout` 파라미터를 0보다 큰 값(초 또는 분)으로 설정합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

예를 들어 `graceful-shutdown-timeout` 값을 `"30m"`으로 늘리면 30분의 제한 시간이지정됩니다. 종료 제한 시간이 종료된 후, 쿼리 작업이 완료될 때까지 대기 중인 경우 서비스 해제 중으로 표시된 노드가 강제로 종료되고 쿼리가 실패합니다. 쿼리 작업이 5분 내에 완료되는 경우, 다른 YARN 애플리케이션이 실행을 완료하면 서비스 해제 중으로 표시된 노드가 5분되는 시점에 종료됩니다.

Example 단계적 서비스 해제가 있는 Presto 자동 조정 구성 예제

`graceful-shutdown-timeout` 값을 설정에 적절한 시간(분)으로 바꿉니다. 최대값이 없습니다. 다음 예제에서는 `1800`초(30분)의 제한 시간 값을 설정합니다.

```
[
    {
        "classification": "presto-config",
        "properties": {
            "graceful-shutdown-timeout": "1800s"
        }
    }
]
```

제한 사항

PrestoDB 단계적 서비스 해제는 `http-server.http.enabled`가 `false`로 설정된 경우와 같이 HTTP 연결이 비활성화된 EMR 클러스터에서는 작동하지 않습니다. Trino는 `http-server.http.enabled` 설정과 관계없이 단계적 서비스 해제를 전혀 지원하지 않습니다.

Amazon EMR 기반 Presto에 대한 고려 사항

Amazon EMR 기반 Presto를 실행할 경우 다음과 같은 제한 사항을 고려합니다.

Presto 명령줄 실행 파일

Amazon EMR에서 PrestoDB 및 Trino는 모두 다음 예제와 같이 동일한 명령줄 실행 파일(`presto-cli`)을 사용합니다.

```
presto-cli --catalog hive
```

구성할 수 없는 Presto 배포 속성

사용하는 Amazon EMR 버전에 따라 사용 가능한 Presto 배포 구성이 결정됩니다. 이러한 구성 속성에 대한 자세한 내용은 Presto 설명서에서 Deploying Presto를 참조하세요. 아래 테이블에는 Presto `properties` 파일의 여러 가지 구성 옵션이 나와 있습니다.

파일	구성 가능
`log.properties`	PrestoDB: Amazon EMR 버전 4.0.0 이상에서 구성할 수 있습니다. `presto-log` 구성 분류를 사용하십시오.
`config.properties`	PrestoDB: Amazon EMR 버전 4.0.0 이상에서 구성할 수 있습니다. `presto-config` 구성 분류를 사용하십시오.
`hive.properties`	PrestoDB: Amazon EMR 버전 4.1.0 이상에서 구성할 수 있습니다. `presto-connector-hive` 구성 분류를 사용하십시오.
`node.properties`	PrestoDB: Amazon EMR 버전 5.6.0 이상에서 구성할 수 있습니다. `presto-node` 구성 분류를 사용하십시오.
`jvm.config`	구성할 수 없습니다.

PrestoDB 설치

애플리케이션 이름(Presto)은 클러스터에 PrestoDB를 설치하는 데 계속 사용됩니다.

PrestoDB 또는 Trino 중 하나를 설치할 수 있지만 단일 클러스터에 둘 다 설치할 수는 없습니다. 클러스터를 생성하려고 할 때 PrestoDB 및 Trino를 모두 지정하면 검증 오류가 발생하고 클러스터 생성 요청이 실패합니다.

EMRFS 및 PrestoS3FileSystem 구성

Amazon EMR 버전 5.12.0 이상에서 PrestoDB는 EMRFS를 사용할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR 파일 시스템(EMRFS)을 참조하세요. 이전 버전의 Amazon EMR에서는 Prestos3 파일 시스템이 유일한 구성 옵션입니다.

보안 구성을 사용하여 Amazon S3의 EMRFS 데이터에 대한 암호화를 설정할 수 있습니다. Amazon S3에 대한 EMRFS 요청에서 IAM 역할을 사용할 수도 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 암호화 옵션 이해 및 Amazon S3에 대한 EMRFS 요청을 위해 IAM 역할 구성을 참조하세요.

Note

Amazon EMR 버전 5.12.0을 사용하여 Amazon S3에서 기본 데이터를 쿼리하는 경우 Presto 오류가 발생할 수 있습니다. 이는 Presto가 `emrfs-site.xml`에서 구성 분류 값을 가져오지 못하기 때문입니다. 이 문제를 해결하려면 `usr/lib/presto/plugin/hive-hadoop2/` 아래에 `emrfs` 하위 디렉터리를 생성하고 `usr/lib/presto/plugin/hive-hadoop2/emrfs`에서 기존 `/usr/share/aws/emr/emrfs/conf/emrfs-site.xml` 파일에 대한 기호 링크를 생성합니다. 그런 다음, presto-server 프로세스를 다시 시작합니다(`sudo presto-server stop` 및 `sudo presto-server start`의 순서대로)

EMRFS 기본값을 재정의하여 PrestoS3FileSystem을 대신 사용해도 됩니다. 그러려면 다음 예제와 같이 `presto-connector-hive` 구성 분류를 사용하여 `hive.s3-file-system-type`을 `PRESTO`로 설정하십시오. 자세한 내용은 [애플리케이션 구성](#) 단원을 참조하십시오.

```
[
   {
      "Classification": "presto-connector-hive",
      "Properties": {
         "hive.s3-file-system-type": "PRESTO"
      }
   }
]
```

PrestoS3FileSystem을 사용한다면 `presto-connector-hive` 구성 분류에 따라 PrestoS3FileSystem 속성을 구성하십시오. 사용 가능한 속성에 대한 자세한 내용은 Presto 설명서의 Hive Connector 섹션에서 [Amazon S3 configuration](#)을 참조하세요. 이러한 설정은 EMRFS에는 적용되지 않습니다.

최종 사용자 위장의 기본 설정

Amazon EMR 버전 5.12.0 이상에서는 HDFS에 액세스할 때 기본적으로 최종 사용자 위장을 활성화합니다. 자세한 내용은 Presto 설명서에서 [End user impersonation](#)을 참조하세요. `presto-config` 구성 분류로 이 설정을 변경하여 `hive.hdfs.impersonation.enabled` 속성을 `false`로 설정합니다.

Presto 웹 인터페이스의 기본 포트

Amazon EMR은 기본적으로 PrestoDB 및 Trino에 대해 Presto 코디네이터의 Presto 웹 인터페이스에서 포트 8889를 사용하도록 구성합니다. 포트를 변경하려면 `presto-config` 구성 분류를 사용하여 `http-server.http.port` 속성을 설정합니다. 자세한 내용은 Presto 설명서에서 Deploying Presto 섹션의 Config properties를 참조하세요.

일부 릴리스의 Hive 버킷 실행 문제

Presto 버전 152.3에는 일부 상황에서 Presto 쿼리 성능이 크게 저하되는 Hive 버킷 실행 문제가 있습니다. Amazon EMR 버전 5.0.3, 5.1.0 및 5.2.0에는 이 버전의 Presto가 포함되어 있습니다. 이 문제를 완화하려면 다음 예제와 같이 `presto-connector-hive` 구성 분류를 사용하여 `hive.bucket-execution` 속성을 `false`로 설정합니다.

```
[
   {
      "Classification": "presto-connector-hive",
      "Properties": {
         "hive.bucket-execution": "false"
      }
   }
]
```

Presto 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Presto의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Presto 버전 정보

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-7.9.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
		kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.8.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.7.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-7.6.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.5.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-7.4.0	0.287	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.3.0	0.285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-7.2.0	0.285	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.36.2	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-7.1.0	0.284	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-7.0.0	0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.15.0	0.283	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.14.0	0.281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.13.0	0.281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.12.0	0.281	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.11.1	0.279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.11.0	0.279	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.10.1	0.278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.10.0	0.278	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.9.1	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.9.0	0.276	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.8.1	0.273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.8.0	0.273	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.7.0	0.272	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.36.1	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.36.0	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.6.0	0.267	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.35.0	0.266	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.5.0	0.261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.4.0	0.254.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.3.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.3.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.2.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.2.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.1.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.1.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-6.0.1	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-6.0.0	0.230	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.34.0	0.261	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.33.1	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.33.0	0.245.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.32.1	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.32.0	0.240.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.31.1	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.31.0	0.238.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.30.2	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.30.1	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.30.0	0.232	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mariadb-server, presto-coordinator, presto-worker
emr-5.29.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.28.1	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.28.0	0.227	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hudi, hudi-presto, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.27.1	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.27.0	0.224	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.26.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.25.0	0.220	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.24.1	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.24.0	0.219	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.23.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.23.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.22.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.21.2	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.21.1	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.21.0	0.215	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.20.1	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.20.0	0.214	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.19.1	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.19.0	0.212	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.18.1	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.18.0	0.210	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.17.2	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.17.1	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.17.0	0.206	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.16.1	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.16.0	0.203	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.15.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.15.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.14.2	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.14.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.14.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.13.1	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.13.0	0.194	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.12.3	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.2	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.12.1	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.12.0	0.188	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.11.4	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.3	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.11.2	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.11.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.11.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.10.1	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.10.0	0.187	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.9.1	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.9.0	0.184	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.8.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.8.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.8.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.7.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.7.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.6.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.6.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.4	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.3	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.5.2	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.1	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker
emr-5.5.0	0.170	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-worker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.4.1	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.4.0	0.166	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.3.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.3.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.3.0	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.2.3	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.2.2	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.2.1	0.157.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.2.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.1.1	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.1.0	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.0.3	0.152.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker

Amazon EMR 릴리스 레이블	Presto 버전	Presto와 함께 설치된 구성 요소
emr-5.0.2	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.0.1	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker
emr-5.0.0	0.150	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hive-client, hcatalog-server, mysql-server, presto-coordinator, presto-wo rker

Apache Spark

Apache Spark는 Amazon EMR 클러스터에서 기계 학습, 스트림 처리 또는 그래픽 분석을 수행하는 데 도움이 되는 분산 처리 프레임워크 및 프로그래밍 모델입니다. Apache 하둡과 마찬가지로, Spark는 빅 데이터 워크로드에 일반적으로 사용되는 오픈 소스 분산형 처리 시스템입니다. 하지만 Spark와 Hadoop MapReduce 간에는 다른 몇 가지 큰 차이가 있습니다. Spark에는 최적화된 비순환 방향 그래프(DAG) 실행 엔진이 있고 메모리 안에 데이터를 능동적으로 캐시하므로 특히 특정 알고리즘 및 대화형 쿼리의 성능이 향상됩니다.

Spark는 기본적으로 Scala, Python 및 Java로 작성된 애플리케이션을 지원합니다. 또한 SQL(Spark), 기계 학습(MLlib), 스트림 처리(Spark 스트리밍) 및 그래프 처리(GraphX)를 위한 몇 가지 긴밀하게 통합된 라이브러리가 포함되어 있습니다. 이러한 도구를 사용하면 다양한 사용 사례에 대해 Spark 프레임워크를 보다 쉽게 활용할 수 있습니다.

Amazon EMR 클러스터에 다른 Hadoop 애플리케이션과 함께 Spark를 설치할 수 있으며, Amazon EMR 파일 시스템(EMRFS)을 활용하여 Amazon S3에서 데이터에 직접 액세스할 수도 있습니다. Hive도 Spark와 통합되므로 HiveContext 객체를 사용하여 Spark에서 Hive 스크립트를 실행할 수 있습니다. Hive 컨텍스트는 spark-shell에 `sqlContext`로 포함됩니다.

Spark로 EMR 클러스터를 설정하고 샘플 데이터 세트를 분석하는 예제 자습서는 AWS 뉴스 블로그의 자습서: Amazon EMR 시작하기를 참조하세요.

> **Important**
>
> Amazon EMR 릴리스 5.16.0부터 사용할 수 있는 Apache Spark 버전 2.3.1은 CVE-2018-8024 및 CVE-2018-1334를 처리합니다. Spark의 이전 버전을 Spark 버전 2.3.1 이상으로 마이그레이션하는 것이 좋습니다.

다음 테이블에는 Amazon EMR이 Spark를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Spark의 버전이 나열되어 있습니다.

이 릴리스에서 Spark와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Spark 버전 정보

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-7.9.0	Spark 3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

다음 테이블에는 Amazon EMR이 Spark를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Spark의 버전이 나열되어 있습니다.

이 릴리스에서 Spark와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Spark 버전 정보

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.15.0	Spark 3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
		nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Note

Amazon EMR 릴리스 6.8.0은 Apache Spark 3.3.0과 함께 제공됩니다. 이번 Spark 릴리스에서는 Apache Log4j 2 및 `log4j2.properties` 파일을 사용하여 Spark 프로세스에서 Log4j를 구성합니다. 클러스터에서 Spark를 사용하거나 사용자 지정 구성 파라미터를 사용하여 EMR 클러스터를 생성하고 Amazon EMR 릴리스 6.8.0으로 업그레이드하려는 경우 Apache Log4j 2의 새로운 `spark-log4j2` 구성 분류 및 키 형식으로 마이그레이션해야 합니다. 자세한 내용은 Apache Log4j 1.x에서 Log4j 2.x로 마이그레이션 단원을 참조하십시오.

다음 테이블에는 Amazon EMR이 Spark를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Spark의 버전이 나열되어 있습니다.

이 릴리스에서 Spark와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Spark 버전 정보

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.36.2	Spark 2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

주제

Apache Spark가 설치된 클러스터 생성

다음 절차에서는 Amazon EMR 콘솔에서 고급 옵션을 사용하여 설치된 Spark로 클러스터를 생성합니다.

고급 옵션을 사용하여 클러스터 설정을 추가적으로 사용자 지정하고 단계를 제출하여 애플리케이션을 프로그래밍 방식으로 설치한 다음, 사용자 지정 애플리케이션을 실행할 수 있습니다. 클러스터 생성 옵션을 사용하여 AWS Glue를 Spark SQL 메타스토어로 사용하도록 선택할 수 있습니다. 자세한 정보는 Amazon EMR에서 Spark와 함께 AWS Glue 데이터 카탈로그 사용을 참조하세요.

Spark가 설치된 클러스터를 시작하려면

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 빠른 옵션을 사용하려면 클러스터 생성을 선택합니다.
3. 클러스터 이름을 입력합니다. 클러스터 이름에는 <, >, $, | 또는 `(백틱) 문자를 포함할 수 없습니다.
4. 소프트웨어 구성에서 릴리스 옵션을 선택합니다.
5. 애플리케이션에서 Spark 애플리케이션 번들을 선택합니다.
6. 필요에 따라 다른 옵션을 선택한 다음 클러스터 생성을 선택합니다.

> **Note**
>
> 클러스터를 생성할 때 Spark를 구성하려면 Spark 구성 단원을 참조하십시오.

를 사용하여 Spark가 설치된 클러스터를 시작하려면 AWS CLI

- 다음 명령을 사용하여 클러스터를 생성합니다.

```
aws emr create-cluster --name "Spark cluster" --release-label emr-7.9.0 --
applications Name=Spark \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 --use-
default-roles
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

SDK for Java를 사용하여 Spark가 설치된 클러스터를 시작하는 방법

`SupportedProductConfig`에 사용된 `RunJobFlowRequest`와 함께 Spark를 애플리케이션으로 지정합니다.

- 다음은 Java를 사용하여 Spark가 설치된 클러스터를 생성하는 방법을 보여주는 예제입니다.

```
import com.amazonaws.AmazonClientException;
import com.amazonaws.auth.AWSCredentials;
import com.amazonaws.auth.AWSStaticCredentialsProvider;
import com.amazonaws.auth.profile.ProfileCredentialsProvider;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.*;
import com.amazonaws.services.elasticmapreduce.util.StepFactory;

public class Main {

        public static void main(String[] args) {
                AWSCredentials credentials_profile = null;
                try {
                        credentials_profile = new
 ProfileCredentialsProvider("default").getCredentials();
                } catch (Exception e) {
                        throw new AmazonClientException(
                                        "Cannot load credentials from .aws/
credentials file. " +
                                                        "Make sure that the
 credentials file exists and the profile name is specified within it.",
                                        e);
                }

                AmazonElasticMapReduce emr =
 AmazonElasticMapReduceClientBuilder.standard()
```

```
                                .withCredentials(new
 AWSStaticCredentialsProvider(credentials_profile))
                                .withRegion(Regions.US_WEST_1)
                                .build();

                // create a step to enable debugging in the AWS Management Console
                StepFactory stepFactory = new StepFactory();
                StepConfig enabledebugging = new StepConfig()
                                .withName("Enable debugging")
                                .withActionOnFailure("TERMINATE_JOB_FLOW")

 .withHadoopJarStep(stepFactory.newEnableDebuggingStep());

                Application spark = new Application().withName("Spark");

                RunJobFlowRequest request = new RunJobFlowRequest()
                                .withName("Spark Cluster")
                                .withReleaseLabel("emr-5.20.0")
                                .withSteps(enabledebugging)
                                .withApplications(spark)
                                .withLogUri("s3://path/to/my/logs/")
                                .withServiceRole("EMR_DefaultRole")
                                .withJobFlowRole("EMR_EC2_DefaultRole")
                                .withInstances(new JobFlowInstancesConfig()
                                                .withEc2SubnetId("subnet-12ab3c45")
                                                .withEc2KeyName("myEc2Key")
                                                .withInstanceCount(3)

 .withKeepJobFlowAliveWhenNoSteps(true)
                                                .withMasterInstanceType("m4.large")

 .withSlaveInstanceType("m4.large"));
                RunJobFlowResult result = emr.runJobFlow(request);
                System.out.println("The cluster ID is " + result.toString());
        }
}
```

Amazon EMR 6.x에서 Docker를 사용하여 Spark 애플리케이션 실행

Amazon EMR 6.0.0을 사용하는 경우 Spark 애플리케이션이 클러스터의 개별 Amazon EC2 인스턴스에 종속 항목을 설치하는 대신 Docker 컨테이너를 사용하여 라이브러리 종속 항목을 정의할 수 있습니

다. 도커로 Spark를 실행하려면 먼저 도커 레지스트리를 구성하고 Spark 애플리케이션을 제출할 때 추가 파라미터를 정의해야 합니다. 자세한 내용은 Docker 통합 구성을 참조하세요.

애플리케이션이 제출되면 YARN이 도커를 호출하여 지정된 도커 이미지를 가져와서 도커 컨테이너 내에서 Spark 애플리케이션을 실행합니다. 이를 통해 종속 항목을 쉽게 정의하고 분리할 수 있습니다. 작업 실행에 필요한 라이브러리를 사용하여 Amazon EMR 클러스터에서 인스턴스를 부트스트랩하거나 준비하는 시간이 단축됩니다.

Docker에서 Spark를 실행할 때의 고려 사항

도커로 Spark를 실행하기 전에 다음 사전 요구 사항을 충족하는지 확인하세요.

- `docker` 패키지와 CLI는 코어 및 작업 노드에만 설치해야 합니다.
- Amazon EMR 6.1.0 이상에서는 다음 명령을 사용하여 프라이머리 노드에 Docker를 설치할 수도 있습니다.
  - ```
    sudo yum install -y docker
    sudo systemctl start docker
    ```
- `spark-submit` 명령은 항상 Amazon EMR 클러스터의 프라이머리 인스턴스에서 실행해야 합니다.
- 클러스터를 시작할 때 추가 파라미터를 정의하는 `container-executor` 분류 키와 함께 분류 API를 사용하여 도커 이미지를 확인하는 데 사용되는 도커 레지스트리를 정의해야 합니다.
  - `docker.trusted.registries`
  - `docker.privileged-containers.registries`
- 도커 컨테이너에서 Spark 애플리케이션을 실행하려면 다음 구성 옵션이 필요합니다.
  - `YARN_CONTAINER_RUNTIME_TYPE=docker`
  - `YARN_CONTAINER_RUNTIME_DOCKER_IMAGE={DOCKER_IMAGE_NAME}`
- Amazon ECR을 사용하여 도커 이미지를 검색하는 경우 클러스터를 자체적으로 인증하도록 구성해야 합니다. 이렇게 하려면 다음 구성 옵션을 사용해야 합니다.
  - YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON
- Amazon EMR 6.1.0 이상에서는 ECR 자동 인증 기능이 활성화된 경우 나열된 명령 `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG={DOCKER_CLIENT_CONFIG_PATH_ON_HDFS` 사용할 필요가 없습니다.
- Spark와 함께 사용되는 모든 도커 이미지에는 Java가 설치되어 있어야 합니다.

필수 조건에 대한 자세한 내용은 Docker 통합 구성을 참조하세요.
    ```

도커 이미지 생성

도커 이미지는 이미지에 포함할 패키지와 구성을 정의하는 Dockerfile을 사용하여 생성됩니다. 다음 두 가지 예제 Dockerfile은 PySpark와 SparkR을 사용합니다.

PySpark Dockerfile

이 Dockerfile에서 생성된 도커 이미지에는 Python 3 및 NumPy Python 패키지가 포함되어 있습니다. 이 Dockerfile은 Amazon Linux 2와 Amazon Corretto JDK 8을 사용합니다.

```
FROM amazoncorretto:8

RUN yum -y update
RUN yum -y install yum-utils
RUN yum -y groupinstall development

RUN yum list python3*
RUN yum -y install python3 python3-dev python3-pip python3-virtualenv

RUN python -V
RUN python3 -V

ENV PYSPARK_DRIVER_PYTHON python3
ENV PYSPARK_PYTHON python3

RUN pip3 install --upgrade pip
RUN pip3 install numpy pandas

RUN python3 -c "import numpy as np"
```

SparkR Dockerfile

이 Dockerfile에서 생성된 도커 이미지에는 R 및 randomForest CRAN 패키지가 포함되어 있습니다. 이 Dockerfile은 Amazon Linux 2와 Amazon Corretto JDK 8을 사용합니다.

```
FROM amazoncorretto:8

RUN java -version

RUN yum -y update
RUN amazon-linux-extras install R4
```

```
RUN yum -y install curl hostname

#setup R configs
RUN echo "r <- getOption('repos'); r['CRAN'] <- 'http://cran.us.r-project.org';
 options(repos = r);" > ~/.Rprofile

RUN Rscript -e "install.packages('randomForest')"
```

Dockerfile 구문에 대한 자세한 내용은 Dockerfile 참조 설명서를 참조하십시오.

Amazon ECR에서 도커 이미지 사용

Amazon Elastic Container Registry(Amazon ECR)는 개발자가 Docker 컨테이너 이미지를 간편하게 저장, 관리 및 배포할 수 있게 해주는 완전관리형 Docker 컨테이너 레지스트리입니다. Amazon ECR을 사용하는 경우 ECR 인스턴스를 신뢰하도록 클러스터를 구성해야 하며 클러스터가 Amazon ECR의 도커 이미지를 사용할 수 있도록 인증을 구성해야 합니다. 자세한 내용은 Amazon ECR에 액세스하도록 YARN 구성을 참조하세요.

Amazon EMR 호스트가 Amazon ECR에 저장된 이미지에 액세스할 수 있도록 하려면 클러스터에 인스턴스 프로파일과 연결된 `AmazonEC2ContainerRegistryReadOnly` 정책의 권한이 있어야 합니다. 자세한 내용은 `AmazonEC2ContainerRegistryReadOnly` 정책을 참조하십시오.

이 예제에서는 Amazon ECR 레지스트리를 신뢰할 수 있도록 다음 추가 구성으로 클러스터를 생성해야 합니다. *`123456789123.dkr.ecr.us-east-1.amazonaws.com`* 엔드포인트를 Amazon ECR 엔드포인트로 바꿉니다.

```
[
  {
    "Classification": "container-executor",
    "Configurations": [
      {
        "Classification": "docker",
        "Properties": {
          "docker.privileged-containers.registries":
 "local,centos,123456789123.dkr.ecr.us-east-1.amazonaws.com",
          "docker.trusted.registries": "local,centos,123456789123.dkr.ecr.us-
east-1.amazonaws.com"
        }
      }
    ],
    "Properties": {}
  }
```

```
]
```

Amazon ECR과 함께 PySpark 사용

다음 예제에서는 PySpark Dockerfile을 사용하며, 이 Dockerfile을 태그 지정하여 Amazon ECR에 업로드합니다. Dockerfile을 업로드한 후 PySpark 작업을 실행하고 Amazon ECR의 도커 이미지를 참조할 수 있습니다.

클러스터를 시작한 후 SSH를 사용하여 코어 노드에 연결하고 다음 명령을 실행하여 PySpark Dockerfile 예제의 로컬 도커 이미지를 빌드합니다.

먼저 디렉터리와 Dockerfile을 만듭니다.

```
mkdir pyspark
vi pyspark/Dockerfile
```

PySpark Dockerfile의 내용을 붙여 넣고 다음 명령을 실행하여 도커 이미지를 빌드합니다.

```
sudo docker build -t local/pyspark-example pyspark/
```

예제용 `emr-docker-examples` ECR 리포지토리를 생성합니다.

```
aws ecr create-repository --repository-name emr-docker-examples
```

로컬로 빌드한 이미지를 태그 지정하고 ECR에 업로드합니다. 이때 *`123456789123.dkr.ecr.us-east-1.amazonaws.com`*을 해당 ECR 엔드포인트로 대체합니다.

```
sudo docker tag local/pyspark-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:pyspark-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:pyspark-example
```

SSH를 사용하여 프라이머리 노드에 연결하고 `main.py`라는 파일 이름으로 Python 스크립트를 준비합니다. 다음 내용을 `main.py` 파일에 붙여 넣고 저장합니다.

```
from pyspark.sql import SparkSession
spark = SparkSession.builder.appName("docker-numpy").getOrCreate()
sc = spark.sparkContext
```

```
import numpy as np
a = np.arange(15).reshape(3, 5)
print(a)
```

Amazon EMR 6.0.0에서 작업을 제출하려면 도커 이미지의 이름을 참조합니다. 추가 구성 파라미터를 정의하여 작업 실행에 Docker가 런타임으로 사용되도록 합니다. Amazon ECR을 사용하는 경우 `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG`에서 Amazon ECR에 대한 인증에 사용되는 보안 인증이 포함된 `config.json` 파일을 참조해야 합니다.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--num-executors 2 \
main.py -v
```

Amazon EMR 6.1.0 이상에서 작업을 제출하려면 도커 이미지의 이름을 참조합니다. ECR 자동 인증이 활성화된 경우 다음 명령을 실행합니다.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:pyspark-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--num-executors 2 \
main.py -v
```

작업이 완료되면 YARN 애플리케이션 ID를 기록하고 다음 명령을 사용하여 PySpark 작업의 출력을 얻습니다.

```
yarn logs --applicationId application_id | grep -C2 '\[\['
LogLength:55
LogContents:
[[ 0  1  2  3  4]
 [ 5  6  7  8  9]
 [10 11 12 13 14]]
```

Amazon ECR과 함께 SparkR 사용

다음 예제에서는 SparkR Dockerfile을 사용하며, 이 Dockerfile을 태그 지정하여 ECR에 업로드합니다. Dockerfile이 업로드된 후 SparkR 작업을 실행하고 Amazon ECR의 도커 이미지를 참조할 수 있습니다.

클러스터를 시작한 후 SSH를 사용하여 코어 노드에 연결하고 다음 명령을 실행하여 SparkR Dockerfile 예제의 로컬 도컬 이미지를 빌드합니다.

먼저 디렉터리와 Dockerfile을 만듭니다.

```
mkdir sparkr
vi sparkr/Dockerfile
```

SparkR Dockerfile의 내용을 붙여 넣고 다음 명령을 실행하여 도커 이미지를 빌드합니다.

```
sudo docker build -t local/sparkr-example sparkr/
```

로컬로 빌드한 이미지를 태그 지정하고 Amazon ECR에 업로드합니다. 이때 *123456789123.dkr.ecr.us-east-1.amazonaws.com*을 해당 Amazon ECR 엔드포인트로 대체합니다.

```
sudo docker tag local/sparkr-example 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:sparkr-example
sudo docker push 123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-examples:sparkr-example
```

SSH를 사용하여 프라이머리 노드에 연결하고 `sparkR.R`이라는 이름으로 R 스크립트를 준비합니다. 다음 내용을 `sparkR.R` 파일에 붙여 넣습니다.

```
library(SparkR)
```

```
sparkR.session(appName = "R with Spark example", sparkConfig =
 list(spark.some.config.option = "some-value"))

sqlContext <- sparkRSQL.init(spark.sparkContext)
library(randomForest)
# check release notes of randomForest
rfNews()

sparkR.session.stop()
```

Amazon EMR 6.0.0에서 작업을 제출하려면 도커 이미지의 이름을 참조합니다. 추가 구성 파라미터를 정의하여 작업 실행에 Docker가 런타임으로 사용되도록 합니다. Amazon ECR을 사용하는 경우 `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG`에서 ECR에 대한 인증에 사용되는 보안 인증이 포함된 `config.json` 파일을 참조해야 합니다.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
DOCKER_CLIENT_CONFIG=hdfs:///user/hadoop/config.json
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG=
$DOCKER_CLIENT_CONFIG \
sparkR.R
```

Amazon EMR 6.1.0 이상에서 작업을 제출하려면 도커 이미지의 이름을 참조합니다. ECR 자동 인증이 활성화된 경우 다음 명령을 실행합니다.

```
DOCKER_IMAGE_NAME=123456789123.dkr.ecr.us-east-1.amazonaws.com/emr-docker-
examples:sparkr-example
spark-submit --master yarn \
--deploy-mode cluster \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.executorEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_TYPE=docker \
--conf spark.yarn.appMasterEnv.YARN_CONTAINER_RUNTIME_DOCKER_IMAGE=$DOCKER_IMAGE_NAME \
sparkR.R
```

작업이 완료되면 YARN 애플리케이션 ID를 기록하고 다음 명령을 사용하여 SparkR 작업의 출력을 얻습니다. 이 예제에는 RandomForest 라이브러리, 설치된 버전 및 릴리스 정보를 사용할 수 있는지 확인하는 테스트가 포함되어 있습니다.

```
yarn logs --applicationId application_id | grep -B4 -A10 "Type rfNews"
randomForest 4.6-14
Type rfNews() to see new features/changes/bug fixes.
Wishlist (formerly TODO):

* Implement the new scheme of handling classwt in classification.

* Use more compact storage of proximity matrix.

* Allow case weights by using the weights in sampling?

========================================================================
Changes in 4.6-14:
```

Amazon EMR에서 Spark와 함께 AWS Glue 데이터 카탈로그 사용

Amazon EMR 릴리스 5.8.0 이상을 사용하면 Glue 데이터 카탈로그를 Apache Hive AWS 메타스토어로 사용하도록 Spark를 구성할 수 있습니다. 영구 Hive 메타스토어 또는 다른 클러스터, 서비스, 애플리케이션 또는 AWS 계정에서 공유하는 Hive 메타스토어가 필요한 경우이 구성을 사용하는 것이 좋습니다.

Amazon EMR 릴리스 6.5.0 이상을 사용하면 Apache Iceberg와 함께 AWS Glue 데이터 카탈로그를 사용하도록 Spark를 구성할 수 있습니다.

Amazon EMR 릴리스 7.5.0 이상을 사용하면 AWS Glue 데이터 카탈로그를 Iceberg REST 카탈로그로 사용하도록 Spark를 구성할 수 있습니다.

AWS Glue는 완전 관리형 추출, 변환 및 로드(ETL) 서비스로, 데이터를 분류하고 정리하고 보강하고 다양한 데이터 스토어 간에 안정적으로 이동할 수 있는 간단하고 비용 효율적입니다. Glue 데이터 카탈로그는 Amazon EMR뿐만 아니라 Amazon RDS, Amazon Redshift, Redshift Spectrum, Athena 및 Apache Hive AWS 메타스토어와 호환되는 모든 애플리케이션과 통합하여 다양한 데이터 소스 및 데이터 형식에 걸쳐 통합된 메타데이터 리포지토리를 제공합니다. AWS Glue 크롤러는 Amazon S3의 소스 데이터에서 스키마를 자동으로 추론하고 관련 메타데이터를 데이터 카탈로그에 저장할 수 있습니다. 데이터 카탈로그에 대한 자세한 내용은 AWS Glue 개발자 안내서의 Glue 데이터 카탈로그 채우기를 참조하세요. AWS

AWS Glue에는 별도의 요금이 적용됩니다. 데이터 카탈로그의 메타데이터 저장 및 액세스에 대한 월별 요금, Glue ETL 작업 및 AWS 크롤러 런타임에 대해 분당 청구되는 시간당 요금, 프로비저닝된 각 개발 엔드포인트에 대해 분당 청구되는 시간당 요금이 있습니다. 데이터 카탈로그에는 최대 100만 개의 객체를 무료로 저장할 수 있습니다. 100만 개 이상의 객체를 저장하는 경우, 100만 이상의 객체 100,000개마다 1USD가 부과됩니다. 데이터 카탈로그의 객체로는 테이블, 파티션 또는 데이터베이스가 있습니다. 자세한 내용은 Glue 요금을 참조하십시오.

> **Important**
>
> 2017년 8월 14일 이전에 Amazon Athena 또는 Amazon Redshift Spectrum을 사용하여 테이블을 생성한 경우 데이터베이스 및 테이블은 AWS Glue 데이터 카탈로그와 별도의 Athena 관리형 카탈로그에 저장됩니다. Amazon EMR을 이러한 테이블과 통합하려면 AWS Glue 데이터 카탈로그로 업그레이드해야 합니다. 자세한 내용은 Amazon Athena 사용 설명서의 AWS Glue 데이터 카탈로그로 업그레이드를 참조하세요.

Glue 데이터 카탈로그를 Apache Hive AWS 메타스토어로 지정

AWS Management Console AWS CLI또는 Amazon EMR API를 사용하여 AWS Glue 데이터 카탈로그를 메타스토어로 지정할 수 있습니다. CLI 또는 API를 사용할 때 Spark에 대한 구성 분류를 사용하여 데이터 카탈로그를 지정합니다. 또한 Amazon EMR 5.16.0 이상에서는 구성 분류를 사용하여 다른에서 데이터 카탈로그를 지정할 수 있습니다 AWS 계정. 콘솔을 사용할 때 고급 옵션 또는 빠른 옵션을 사용하여 데이터 카탈로그를 지정할 수 있습니다.

> **Note**
>
> Zeppelin은 Spark 구성 요소와 함께 설치되므로 Zeppelin에서도 AWS Glue 데이터 카탈로그를 사용하는 옵션을 사용할 수 있습니다.

Console

새 콘솔을 사용하여 AWS Glue 데이터 카탈로그를 Apache Hive 메타스토어로 지정하려면

1. 에 로그인 AWS Management Console하고 https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 왼쪽 탐색 창의 Amazon EMR on EC2에서 클러스터를 선택하고 클러스터 생성을 선택합니다.

3. 애플리케이션 번들에서 Spark 또는 사용자 지정을 선택합니다. 클러스터를 사용자 지정하는 경우 Zeppelin 또는 Spark를 애플리케이션 중 하나로 선택해야 합니다.
4. AWS Glue 데이터 카탈로그 설정에서 Spark 테이블 메타데이터에 사용 확인란을 선택합니다.
5. 클러스터에 적용할 다른 옵션을 선택합니다.
6. 클러스터를 시작하려면 클러스터 생성을 선택합니다.

AWS CLI

를 사용하여 AWS Glue 데이터 카탈로그를 Apache Hive 메타스토어로 지정하려면 AWS CLI

AWS CLI 및 Amazon EMR API를 사용하여 구성 분류를 지정하는 방법에 대한 자세한 내용은 섹션을 참조하세요애플리케이션 구성.

- 다음 예에서 나온 것처럼 `spark-hive-site` 분류를 사용하여 `hive.metastore.client.factory.class`에 대한 값을 지정합니다.

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory"
    }
  }
]
```

다른 AWS 계정에서 데이터 카탈로그를 지정하려면 다음 예제와 같이 `hive.metastore.glue.catalogid` 속성을 추가합니다. *`acct-id`*를 데이터 카탈로그의 AWS 계정으로 바꿉니다.

```
[
  {
    "Classification": "spark-hive-site",
    "Properties": {
      "hive.metastore.client.factory.class":
 "com.amazonaws.glue.catalog.metastore.AWSGlueDataCatalogHiveClientFactory",
      "hive.metastore.glue.catalogid": "acct-id"
    }
  }
```

```
]
```

AWS Glue 데이터 카탈로그를 Apache Iceberg 카탈로그로 지정

AWS 또는 AWS Management Console AWS CLI Amazon EMR API를 사용하거나 Spark 세션 런타임 구성에서 Glue 데이터 카탈로그를 Apache Iceberg 카탈로그 구현 또는 Apache Iceberg REST 카탈로그 엔드포인트로 지정할 수 있습니다. CLI 또는 API를 사용할 때 Spark에 대한 구성 분류를 사용하여 데이터 카탈로그를 지정합니다. 자세한 내용은 AWS Glue 데이터 카탈로그를 Apache Iceberg 카탈로그로 지정을 참조하세요.

IAM 권한

클러스터의 EC2 인스턴스 프로파일에는 AWS Glue 작업에 대한 IAM 권한이 있어야 합니다. 또한 AWS Glue Data Catalog 객체에 대한 암호화를 활성화하는 경우 역할은 암호화에 AWS KMS key 사용되는를 암호화, 복호화 및 생성할 수 있어야 합니다.

AWS Glue 작업에 대한 권한

Amazon EMR에 대한 기본 EC2 인스턴스 프로파일을 사용할 경우 어떤 작업도 필요하지 않습니다. 에 연결된 `AmazonElasticMapReduceforEC2Role` 관리형 정책은 필요한 모든 AWS Glue 작업을 `EMR_EC2_DefaultRole` 허용합니다. 그러나 사용자 지정 EC2 인스턴스 프로파일 및 권한을 지정하는 경우 적절한 AWS Glue 작업을 구성해야 합니다. `AmazonElasticMapReduceforEC2Role` 관리형 정책을 시작점으로 사용합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터 EC2 인스턴스의 서비스 역할(EC2 인스턴스 프로파일)을 참조하세요.

AWS Glue 데이터 카탈로그 암호화 및 복호화 권한

인스턴스 프로파일에 키를 사용하여 데이터를 암호화하고 복호화할 수 있는 권한이 필요합니다. 다음 두 설명이 모두 적용되는 경우에는 이러한 권한을 구성할 필요가 없습니다.

- AWS Glue용 관리형 키를 사용하여 AWS Glue Data Catalog 객체에 대한 암호화를 활성화합니다.
- Glue 데이터 카탈로그 AWS 계정 와 동일한 AWS 에 있는 클러스터를 사용합니다.

그렇지 않으면 EC2 인스턴스 프로파일에 연결된 권한 정책에 다음 명령을 추가해야 합니다.

```
[
    {
```

```
        "Version": "2012-10-17",
        "Statement": [
            {
                "Effect": "Allow",
                "Action": [
                    "kms:Decrypt",
                    "kms:Encrypt",
                    "kms:GenerateDataKey"
                ],
                "Resource": "arn:aws:kms:region:acct-
id:key/12345678-1234-1234-1234-123456789012"
            }
        ]
    }
]
```

Glue 데이터 카탈로그 암호화에 대한 자세한 내용은 AWS AWS Glue 개발자 안내서의 데이터 카탈로그 암호화를 참조하세요.

리소스 기반 권한

Amazon EMR에서 AWS Glue를 Hive, Spark 또는 Presto와 함께 사용하는 경우 AWS Glue는 리소스 기반 정책을 지원하여 데이터 카탈로그 리소스에 대한 액세스를 제어합니다. 이러한 리소스에는 데이터베이스, 테이블, 연결 및 사용자 정의 기능이 포함됩니다. 자세한 내용은 AWS Glue 개발자 안내서에서 AWS Glue 리소스 정책을 참조하세요.

리소스 기반 정책을 사용하여 Amazon EMR 내에서 AWS Glue에 대한 액세스를 제한하는 경우 권한 정책에서 지정하는 보안 주체는 클러스터 생성 시 지정된 EC2 인스턴스 프로파일과 연결된 역할 ARN이어야 합니다. 예를 들어, 카탈로그에 연결된 리소스 기반 정책의 경우 다음 예에 표시된 형식을 사용하여 클러스터 EC2 인스턴스에 대한 기본 서비스 역할의 ARN(*EMR_EC2_DefaultRole*)을 `Principal`로 지정할 수 있습니다.

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*acct-id*는 AWS Glue 계정 ID와 다를 수 있습니다. 그러면 다른 계정의 EMR 클러스터에서 액세스할 수 있습니다. 각각 다른 계정에서 여러 보안 주체를 지정할 수 있습니다.

AWS Glue 데이터 카탈로그 사용 시 고려 사항

Glue Data Catalog를 Spark에서 Apache Hive AWS 메타스토어로 사용할 때는 다음 항목을 고려하세요.

- 테이블을 생성할 때 위치 URI 없이 기본 데이터베이스를 지정하면 오류가 발생합니다. 차선책으로 `LOCATION`을 사용할 때 `s3://amzn-s3-demo-bucket1` 절을 사용하여 버킷 위치를 지정합니다(예: `CREATE TABLE`). 또는 기본 데이터베이스가 아닌 데이터베이스에서 테이블을 생성합니다.
- AWS Glue 내에서 테이블 이름을 바꾸는 것은 지원되지 않습니다.
- `LOCATION`을 지정하지 않고 Hive 테이블을 생성하면 테이블 데이터가 `hive.metastore.warehouse.dir` 속성을 통해 지정된 위치에 저장됩니다. 기본적으로 HDFS에 있는 위치입니다. 다른 클러스터가 테이블에 액세스해야 하는 경우, 테이블을 생성한 클러스터에 대한 적절한 권한이 없으면 실패합니다. 또한 HDFS 스토리지는 일시적이기 때문에, 클러스터가 종료되는 경우 테이블 데이터가 손실되고 테이블을 다시 생성해야 합니다. Glue를 사용하여 Hive 테이블을 생성할 때 Amazon S3`LOCATION`에서 AWS 를 지정하는 것이 좋습니다. 또는 `hive-site` 구성 분류를 사용하여 모든 Hive 테이블에 적용되는 Amazon S3에서 `hive.metastore.warehouse.dir`에 대한 위치를 지정할 수도 있습니다. HDFS 위치에 테이블이 생성되고 테이블을 생성한 클러스터가 여전히 실행 중인 경우 AWS Glue 내에서 테이블 위치를 Amazon S3로 업데이트할 수 있습니다. 자세한 내용은 Glue 개발자 안내서의 AWS Glue 콘솔에서 테이블 작업을 참조하세요. AWS
- 따옴표와 아포스트로피가 포함된 파티션 값은 지원되지 않습니다(예: `PARTITION (owner="Doe's").`).
- 열 통계는 emr-5.31.0 이상에서 지원됩니다.
- Hive 권한 부여 사용은 지원되지 않습니다. 대안으로 AWS Glue 리소스 기반 정책 사용을 고려합니다. 자세한 내용은 AWS Glue 데이터 카탈로그에 대한 Amazon EMR 액세스를 위한 리소스 기반 정책 사용을 참조하세요.

Spark에서 AWS Glue 데이터 카탈로그를 Apache Iceberg REST 카탈로그로 사용할 때는 다음 사항을 고려하세요.

- 에 설명된 Iceberg와 함께 Spark 세션 카탈로그를 사용하는 경우 Iceberg SparkCatalog를 사용할 때와 SparkSessionCatalog를 사용할 때의 구성 차이 AWS Glue 데이터 카탈로그를 Apache Iceberg REST 카탈로그로 구성하는 것 외에도 Glue 데이터 카탈로그를 Apache Hive AWS 메타스토어로 구성해야 합니다.
- AWS Glue Data Catalog IRC 엔드포인트는 Amazon SigV4 인증 체계만 지원합니다. OAuth는 지원되지 않습니다. OAuth 사용자의 경우 IAM Identity Center를 사용하여 액세스를 구성하십시오. Lake Formation과 IAM Identity Center 연결을 참조하세요.
- AWS Glue Iceberg REST 카탈로그는 오픈 소스의 모든 작업을 지원하지 않습니다.

Amazon EMR에서 Spark를 사용하여 AWS Glue Data Catalog의 다중 카탈로그 계층 구조 작업

Amazon EMR 클러스터를 등록하여 AWS Glue Data Catalog에 액세스할 수 있습니다. 그러면 테이블 및 기타 카탈로그 리소스를 다양한 소비자가 사용할 수 있습니다. AWS Glue Data Catalog는 Amazon S3 데이터 레이크에서 데이터를 통합하는 다중 카탈로그 계층 구조를 지원합니다. 또한 데이터에 액세스하기 위한 Hive 메타스토어 API와 오픈 소스 Apache Iceberg REST API를 모두 제공합니다. 이러한 기능은 Amazon EMR 및 Amazon Athena 및 Amazon Redshift와 같은 기타 서비스에서 사용할 수 있습니다.

카탈로그 리소스 구성 방법

Glue 데이터 카탈로그에서 리소스를 생성할 때 Apache Iceberg REST API 또는 Hive AWS 메타스토어를 지원하는 모든 SQL 엔진에서 리소스에 액세스할 수 있습니다. AWS Lake Formation은 권한을 관리합니다.

AWS Glue Data Catalog에서 데이터는 카탈로그, 데이터베이스 및 테이블의 논리적 계층 구조로 구성됩니다.

- 카탈로그 - 스키마 또는 테이블과 같은 데이터 스토어의 객체를 보관하는 논리적 컨테이너입니다.
- Redshift Managed Storage(RMS) 테이블을 저장하는 카탈로그 - 카탈로그를 관리하여 RMS 테이블을 저장하는 경우 Iceberg를 사용하여 이러한 테이블에 액세스할 수 있습니다.
- 데이터베이스 - 카탈로그의 테이블 및 뷰와 같은 데이터 객체를 구성합니다.
- 테이블 및 뷰 - 추상화 계층에 이해하기 쉬운 스키마를 제공하는 데이터베이스의 데이터 객체입니다. 다양한 형식과 다양한 위치에 있을 수 있는 기본 데이터에 액세스할 수 있는 계층을 제공합니다.

Amazon EMR과 함께 사용할 데이터 카탈로그 구성

시작하려면 Amazon EMR 도구를 지원하도록 카탈로그를 구성합니다. AWS Glue 데이터 카탈로그는 Hive 메타스토어 호환성과 Iceberg REST 호환 APIs 제공합니다.

Hive 메타스토어를 사용하여 Amazon EMR 구성

이를 설정하는 방법에 대한 자세한 내용은 AWS Glue 사용 설명서의 Spark 작업에 대한 Glue 데이터 카탈로그 지원을 참조하세요. AWS 이 주제에서는 AWS Glue 데이터 카탈로그를 Hive 메타스토어로 구성하고 엔드포인트로 사용할 수 있도록 하는 방법을 설명합니다. 또한 Glue 데이터 카탈로그를

Spark AWS 용 Apache Hive 메타스토어로 사용에서 Glue 데이터 카탈로그를 Spark AWS 메타스토어로 지정하는 방법을 보여주는 Amazon EMR 설명서가 제공됩니다.

AWS Glue 데이터 카탈로그의 리소스에 액세스하기 위한 권한

이 섹션에서는 카탈로그 데이터와 함께 Amazon EMR 도구를 사용하기 위한 IAM 정책 요구 사항을 설명합니다. 클러스터를 AWS Glue 데이터 카탈로그에 등록한 후 이후에 생성된 데이터 카탈로그의 생성 및 변경 사항을 검색하려면 다음 권한이 필요합니다.

- glue:GetCatalog
- glue:GetCatalogs
- sts:AssumeRole
- sts:TagSession
- sts:SetContext
- sts:SetSourceIdentity

대부분의 경우 권한을 할당할 때 IAM 역할을 생성하고 해당 역할에 권한을 할당하는 것이 좋습니다.

또한 카탈로그 데이터를 쿼리하려면를 사용하여 데이터 카탈로그에 대한 권한을 설정해야 합니다 AWS Lake Formation. 에서 데이터 카탈로그에 대한 권한을 설정하는 방법에 대한 자세한 내용은 데이터 카탈로그 리소스에 대한 권한 부여 및 취소를 AWS Lake Formation참조하세요. https://docs.aws.amazon.com/lake-formation/latest/dg/granting-catalog-permissions.html

클러스터를 생성 및 구성하고 카탈로그 객체에 대한 권한을 설정한 후 작업을 제출하여 데이터를 쿼리하고 처리할 수 있습니다.

AWS Glue 데이터 카탈로그의 다중 카탈로그 계층 구조에 액세스하도록 Spark 구성

EMR 7.5를 사용하면 AWS Glue의 다중 카탈로그 계층 구조를 사용하도록 Spark를 구성할 수 있습니다. 다중 카탈로그 계층 구조를 사용하면 다음을 수행할 수 있습니다.

- 기존 Amazon Redshift 데이터 웨어하우스의 테이블, 뷰 및 구체화된 뷰와 같은 Redshift Managed Storage(RMS) 데이터를 AWS Glue Data Catalog로 가져옵니다. EC2 및 EMR Serverless의 EMR을 사용하여 이러한 객체를 쿼리할 수 있습니다.
- RMS 카탈로그, AWS Glue Data Catalog를 생성하고 ZeroETL을 사용하여 RMS에 데이터를 저장하고 Iceberg 호환 쿼리 엔진을 사용하여 데이터를 쿼리합니다.

- 압축, 스냅샷 및 보존을 포함하는 모든 기능을 갖춘 스토리지 관리를 사용하여 AWS Glue Data Catalog에서 관리형 Iceberg 테이블을 생성합니다.

Spark 세션을 초기화할 때 다중 카탈로그에 연결

다음 예제에서는 대화형 Spark 셸, Spark 제출 또는 Amazon EMR 노트북을 사용하여 AWS Glue의 다중 카탈로그 계층 구조로 작업하는 방법을 보여줍니다.

spark-shell

1. SSH를 사용하여 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. Spark 셸을 시작하려면 다음 명령을 입력합니다. PySpark 쉘을 사용하려면 `spark-shell`을 `pyspark`로 바꿉니다.

```
spark-shell \
    --conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
    --conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket/prefix/
    --conf spark.sql.catalog.my_catalog.type=glue \
    --conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
    --conf spark.sql.defaultCatalog=my_catalog \
    --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

spark-submit

1. SSH를 사용하여 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. 다음 명령을 입력하여 Spark용 Spark 세션을 시작합니다.

```
spark-submit \
--conf spark.sql.catalog.my_catalog=org.apache.iceberg.spark.SparkCatalog \
--conf spark.sql.catalog.my_catalog.warehouse=s3://amzn-s3-demo-bucket1/prefix \
--conf spark.sql.catalog.my_catalog.type=glue \
--conf spark.sql.catalog.my_catalog.glue.id=Glue RMS catalog ID \
--conf spark.sql.defaultCatalog=my_catalog \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

EMR Studio notebooks

EMR Studio 노트북을 사용하여 Spark 세션을 초기화하려면 다음 예제와 같이 Amazon EMR Notebooks에서 `%%configure` 매직 명령을 사용하여 Spark 세션을 구성합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 EMR Notebooks 매직 사용을 참조하세요.

```
%%configure -f{
"conf":{
    "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
    "spark.sql.catalog.my_catalog.type":"glue",
    "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
    "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
    "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
}
```

CLI

CLI를 사용하여 Spark 세션을 초기화하려면 다음 샘플을 실행합니다. AWS CLI 및 Amazon EMR API를 사용하여 구성 분류를 지정하는 방법에 대한 자세한 내용은 애플리케이션 구성을 참조하세요.

```
[
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.catalog.my_catalog":"org.apache.iceberg.spark.SparkCatalog",
      "spark.sql.catalog.my_catalog.type":"glue",
      "spark.sql.catalog.my_catalog.glue.id":"Glue RMS catalog ID",
      "spark.sql.catalog.my_catalog.warehouse":"s3://amzn-s3-demo-bucket1/prefix/",
      "spark.sql.defaultCatalog", "my_catalog",

 "spark.sql.extensions":"org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions"
    }
  }
]
```

AWS Glue 데이터 카탈로그를 사용하여 Redshift 관리형 스토리지로 Spark 세션 초기화

다음 샘플 명령은 AWS Glue 데이터 카탈로그를 사용하여 Spark 세션을 초기화합니다.

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
  --conf spark.sql.catalog.rms.type=glue \
  --conf spark.sql.catalog.rms.glue.id=Glue RMS catalog ID \
  --conf spark.sql.defaultCatalog=rms
  --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

다음 샘플은 Iceberg REST API 및 Redshift Managed Storage with AWS Glue Data Catalog를 사용하여 Spark 세션을 초기화합니다.

```
spark-sql \
  --conf spark.sql.catalog.rms=org.apache.iceberg.spark.SparkCatalog \
  --conf spark.sql.catalog.rms.type=rest \
  --conf spark.sql.catalog.rms.warehouse=glue RMS catalog ID \
  --conf spark.sql.catalog.rms.uri=Glue endpoint URI/iceberg \
  --conf spark.sql.catalog.rms.rest.sigv4-enabled=true \
  --conf spark.sql.catalog.rms.rest.signing-name=glue \
  --conf spark.sql.defaultCatalog=rms \
  --conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions
```

Spark Iceberg에서 AWS Glue 다중 카탈로그 계층 구조를 사용하는 방법에 대한 자세한 내용은 [Spark에서 Iceberg 클러스터 사용을 참조하세요]().

다중 카탈로그 구성에 대한 고려 사항 및 제한 사항

- Apache Hive 메타스토어와 함께 다중 카탈로그 계층 구조를 사용하는 것은 지원되지 않습니다.
- Apache Iceberg와 함께 다중 카탈로그 계층 구조를 사용하면를 사용할 때 Apache Hive 메타스토어로의 폴백을 지원할 수 없습니다`SparkSessionCatalog`.
- 런타임 역할이 있는 EMR on EC2 클러스터는 다중 카탈로그 계층 구조를 지원하지 않습니다.
- 에서 활성화된 EMR on EC2 클러스터는 다중 카탈로그 계층 구조를 지원하지 AWS Lake Formation 않습니다.

Spark 구성

구성 분류를 사용하여 Spark on Amazon EMR을 구성할 수 있습니다. 구성 분류에 대한 자세한 내용은 애플리케이션 구성 섹션을 참조하세요.

Amazon EMR에서 Spark에 대한 구성 분류에는 다음이 포함됩니다.

- **spark** - maximizeResourceAllocation 속성을 true 또는 false로 설정합니다. true로 설정하면, Amazon EMR에서 클러스터 하드웨어 구성을 기반으로 spark-defaults 속성을 자동으로 구성합니다. 자세한 내용은 maximizeResourceAllocation 사용하기 단원을 참조하십시오.
- **spark-defaults** - spark-defaults.conf 파일에서 값을 설정합니다. 자세한 내용은 Spark 설명서에서 Spark configuration을 참조하세요.
- **spark-env** - spark-env.sh 파일에서 값을 설정합니다. 자세한 내용은 Spark 설명서에서 Environment variables를 참조하세요.
- **spark-hive-site** - hive-site.xml에 Spark에 대한 값을 설정합니다.
- **spark-log4j** - (Amazon EMR 릴리스 6.7.x 이하) log4j.properties 파일에서 값을 설정합니다. 자세한 내용은 Github에서 log4j.properties.template 파일을 참조하세요.
- **spark-log4j2** - (Amazon EMR 릴리스 6.8.0 이상) log4j2.properties 파일에서 값을 설정합니다. 자세한 내용은 Github에서 log4j.properties.template 파일을 참조하세요.
- **spark-metrics** - metrics.properties 파일에서 값을 설정합니다. 설정 및 자세한 내용은 Github의 metrics.properties.template 파일 및 Spark 설명서의 Metrics(지표)를 참조하십시오.

> **Note**
>
> 다른 플랫폼에서 Spark 워크로드를 Amazon EMR로 마이그레이션하는 경우 사용자 지정 구성을 추가하기 전에 Amazon EMR에서 설정한 Spark 기본값 을 사용하여 워크로드를 테스트하는 것이 좋습니다. 대부분의 고객은 기본 설정을 통해 성능을 개선할 수 있습니다.

주제

Amazon EMR에서 설정한 Spark 기본값

다음 테이블에서는 Amazon EMR이 애플리케이션에 영향을 미치는 `spark-defaults`의 기본값을 설정하는 방법을 보여줍니다.

Amazon EMR에서 설정한 Spark 기본값

설정	설명	기본값
`spark.executor.memory`	실행기 프로세스당 사용할 메모리 양. 예시: 1g, 2g.	이 설정은 클러스터의 코어 및 태스크 인스턴스 유형에 따라 결정됩니다.
`spark.executor.cores`	각 실행기에서 사용할 코어 수입니다.	이 설정은 클러스터의 코어 및 태스크 인스턴스 유형에 따라 결정됩니다.
`spark.dynamicAllocation.enabled`	true인 경우 동적 리소스 할당을 사용하여 워크로드를 기반으로 애플리케이션에 등록된 실행기 수를 늘리거나 줄입니다.	`true`(Amazon EMR 4.4.0 이상) **Note** Amazon EMR에서 Spark Shuffle Service를 자동으로 구성합니다.
`spark.sql.hive.advancedPartitionPredicatePushdown.enabled`	true인 경우 Hive 메타스토어로의 고급 파티션 조건부 푸시다운이 활성화됩니다.	`true`
`spark.sql.hive.stringLikePartitionPredicatePushdown.enabled`	`startsWith` , `contains`, `endsWith` 필터를 Hive 메타스토어로 푸시다운합니다.	`true`

설정	설명	기본값
	Note Glue는 startsWith , contains 또는 endsWith에 대한 조건부 푸시다운을 지원하지 않습니다. Glue 메타스토어를 사용하는 중에 이러한 함수의 조건부 푸시다운으로 인해 오류가 발생하는 경우 이 구성을 false로 설정합니다.	

Amazon EMR 6.1.0에서 Spark 가비지 수집 구성

spark.driver.extraJavaOptions 및 spark.executor.extraJavaOptions에서 사용자 지정 가비지 수집 구성을 설정하면 Amazon EMR 6.1에서 드라이버 또는 실행기가 시작되지 않습니다. Amazon EMR 6.1.0과 가비지 수집 구성이 충돌하기 때문입니다. Amazon EMR 6.1.0의 경우 기본 가비지 수집 구성은 spark.driver.defaultJavaOptions 및 spark.executor.defaultJavaOptions를 통해 설정됩니다. 이 구성은 Amazon EMR 6.1.0에만 적용됩니다. 로깅 구성과 같이 가비지 수집과 관련이 없는 JVM 옵션(-verbose:class)은 계속 extraJavaOptions를 통해 설정할 수 있습니다. 자세한 내용은 Spark application properties를 참조하세요.

maximizeResourceAllocation 사용하기

클러스터의 각 노드에서 가능한 최대 리소스를 사용하도록 실행기를 구성하려면 spark 구성 분류에서 maximizeResourceAllocation을 true로 설정합니다. maximizeResourceAllocation은 Amazon EMR에만 해당됩니다. maximizeResourceAllocation을 활성화하는 경우 Amazon EMR은 코어 인스턴스 그룹의 인스턴스에 있는 실행기에 대해 사용할 수 있는 최대 컴퓨팅 및 메모리 리소스를 계산합니다. 그런 다음 계산된 최대값을 기반으로 해당 spark-defaults 설정을 설정합니다.

Amazon EMR은 코어 인스턴스 플릿의 인스턴스 유형에 기반하여 실행기에 대해 사용할 수 있는 최대 컴퓨팅 및 메모리 리소스를 계산합니다. 각 인스턴스 플릿은 플릿 내에서 서로 다른 인스턴스 유형과 크기를 가질 수 있으므로 Amazon EMR이 사용하는 실행기 구성이 클러스터에 적합하지 않을 수 있습니다. 따라서 최대 리소스 할당을 사용할 때 기본 설정을 사용하지 않는 것이 좋습니다. 인스턴스 플릿 클러스터에 대한 사용자 지정 설정을 구성합니다.

> **Note**
>
> HBase와 같은 다른 분산 애플리케이션이 있는 클러스터에서는 `maximizeResourceAllocation` 옵션을 사용하지 않아야 합니다. Amazon EMR은 분산 애플리케이션에 사용자 지정 YARN 구성을 사용하므로 `maximizeResourceAllocation`과 충돌하여 Spark 애플리케이션이 실패할 수 있습니다.

다음은 `maximizeResourceAllocation`을 `true`로 설정한 Spark 구성 분류 예제입니다.

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
    }
  }
]
```

`spark-defaults`활성화 시 `maximizeResourceAllocation` 에서 구성되는 설정

설정	설명	값
spark.default.parallelism	사용자가 설정하지 않은 경우 조인, reduceByKey 및 병렬 처리 같은 변환에서 반환한 RDD의 기본 파티션 수입니다.	YARN 컨테이너에서 사용할 수 있는 CPU 코어의 2배수입니다.
spark.driver.memory	SparkContext가 초기화된 경우와 같이 드라이버 프로세스에 사용할 메모리 양(예: 1g, 2g)입니다.	클러스터의 인스턴스 유형을 기반으로 설정이 구성됩니다. 하지만 Spark 드라이버 애플리케이션을 프라이머리 또는 코어 인스턴스 중 하나에서(예를

설정	설명	값
		들면 각각 YARN 클라이언트 및 클러스터 모드에서) 실행할 수 있으므로 설정이 이러한 두 인스턴스 그룹에서 더 작은 인스턴스 유형에 따라 설정됩니다.
spark.executor.memory	실행기 프로세스당 사용하는 메모리 양(예: 1g, 2g)입니다.	클러스터의 코어 및 작업 인스턴스 유형을 기반으로 설정이 구성됩니다.
spark.executor.cores	각 실행기에서 사용할 코어 수입니다.	클러스터의 코어 및 작업 인스턴스 유형을 기반으로 설정이 구성됩니다.
spark.executor.instances	실행기의 수입니다.	클러스터의 코어 및 작업 인스턴스 유형을 기반으로 설정이 구성됩니다. `spark.dynamicAllocation.enabled` 가 동시에 명시적으로 true로 설정되지 않은 경우 설정합니다.

노드 서비스 해제 동작 구성

Amazon EMR 릴리스 5.9.0 이상에서는 Amazon EMR의 Spark에는 수동 크기 조정 또는 자동 조정 정책 요청으로 인해 Spark가 노드 종료를 원활하게 처리하도록 돕는 일련의 기능이 포함되어 있습니다. Amazon EMR은 YARN의 폐기 메커니즘을 기반으로 하여 구축된 Spark에서 블랙리스팅 메커니즘을 구현합니다. 이 메커니즘은 폐기 중인 노드에서 새 작업을 예약하지 않는 동시에 이미 실행 중인 작업을 완료하는 데 도움이 됩니다. 또한 노드가 종료될 때 셔플 블록이 손실되는 경우, Spark 작업을 보다 빠르게 복구하는 데 도움이 되는 기능들도 있습니다. 재계산 프로세스는 더 빨리 트리거되고 단계 재시도 횟수를 줄이면서 더 빨리 재계산하도록 최적화되며, 셔플 블록 누락으로 인한 가져오기 실패로 인해 작업이 실패하는 것을 방지할 수 있습니다.

Important

스팟 인스턴스를 사용할 때 Spark 복원력을 높일 수 있도록 Amazon EMR 릴리스 버전 5.11.0에 `spark.decommissioning.timeout.threshold` 설정이 추가되었습니다. 이전 릴리스에서는 노드가 스팟 인스턴스를 사용하고 입찰 가격으로 인해 인스턴스가 종료될 때 Spark가 정상적으로 종료를 처리하지 못할 수 있습니다. 작업은 실패할 수 있으며 셔플 재계산 시 상당한 시간이 걸릴 수 있습니다. 이러한 이유로 스팟 인스턴스를 사용하는 경우, 5.11.0 이상의 릴리스를 사용하는 것이 좋습니다.

Spark 노드 서비스 해제 설정

설정	설명	기본값
`spark.blacklist.decommissioning.enabled`	`true`로 설정한 경우 Spark는 YARN에서 `decommissioning` 상태의 노드를 거부 목록에 등록합니다. Spark은 해당 노드에서 실행 중인 executor에 대해 새 작업을 예약하지 않습니다. 이미 실행 중인 작업은 완료 할 수 있습니다.	`true`
`spark.blacklist.decommissioning.timeout`	`decommissioning` 상태의 노드가 거부 목록에 등록된 동안의 시간입니다. 기본적으로 이 값은 1시간으로 설정되는데, 이는 `yarn.resourcemanager.decommissioning.timeout` 에 대한 기본값이기도 합니다. 전체 서비스 해제 기간에 노드를 거부 목록에 등록하려면 이 값을 `yarn.resourcemanager.decommissioning.timeout` 이상으로 설정합니다. 서비스 해제 제한 시간이 만료되면 노드는 `decommissioned` 상태로	`1h`

설정	설명	기본값
	전환되며 Amazon EMR은 노드의 EC2 인스턴스를 종료할 수 있습니다. 제한 시간이 만료된 후에도 작업이 계속 실행 중이면 다른 노드에서 실행 중인 executor에서 손실되거나 종료되며 일정이 조정됩니다.	
`spark.decommissioning.timeout.threshold`	Amazon EMR 릴리스 5.11.0 이상에서 사용할 수 있습니다. 초 단위로 지정됩니다. 노드가 서비스 해제 상태로 전환될 때 호스트가 이 값보다 작거나 같은 시간 안에 서비스 해제되는 경우, Amazon EMR은 노드가 서비스 해제 상태로 전환되기를 기다리지 않고 노드를 거부 목록에 등록할 뿐 아니라 호스트 상태를 정리합니다(`spark.resourceManager.cleanupExpiredHost` 에서 지정한 대로). 이를 통해 Spark가 스팟 인스턴스 종료를 더 잘 처리할 수 있습니다. 다른 노드가 셔플 파일을 읽기에는 시간이 부족할 수 있는 `yarn.resourcemager.decommissioning.timeout` 값에 상관없이 20초 제한 시간 이내에 스팟 인스턴스가 폐기되기 때문입니다.	`20s`

설정	설명	기본값
spark.resourceManager.cleanupExpiredHost	true로 설정된 Spark는 decommissioned 상태에 있는 노드의 executor에 저장된 모든 캐시된 데이터와 셔플 블록의 등록을 취소합니다. 이렇게 하면 복구 프로세스가 빨라집니다.	true
spark.stage.attempt.ignoreOnDecommissionFetchFailure	true로 설정되면 폐기된 노드에서 가져오기 실패 건수가 너무 많은 이유로 Spark에서 단계 및 작업이 실패하는 문제를 방지하는 데 도움이 됩니다. decommissioned 상태의 노드에서 셔플 블록 가져오기 실패는 연속 가져오기의 최대 실패 횟수에 포함되지 않습니다.	true

Spark ThriftServer 환경 변수

Spark에서 Hive Thrift Server Port 환경 변수 `HIVE_SERVER2_THRIFT_PORT`를 10001로 설정합니다.

Spark 기본 설정 변경

`spark-defaults` 구성 분류 또는 `spark` 구성 분류의 `maximizeResourceAllocation` 설정을 사용하여 `spark-defaults.conf`에서 기본값을 변경합니다.

다음 절차에는 CLI 또는 콘솔을 사용하여 설정을 수정하는 방법이 나와 있습니다.

CLI를 사용하여 spark.executor.memory가 2g로 설정된 클러스터를 생성하는 방법

- Amazon S3에 저장된 `myConfig.json` 파일을 참조하는 다음 명령을 사용하여 Spark가 설치되고 `spark.executor.memory`가 2g로 설정된 클러스터를 생성합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
```

```
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`myConfig.json`:

```
[
    {
      "Classification": "spark-defaults",
      "Properties": {
        "spark.executor.memory": "2G"
      }
    }
  ]
```

콘솔을 사용하여 spark.executor.memory가 2ㅎ로 설정된 클러스터를 생성하는 방법

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. Spark를 선택합니다.
4. 소프트웨어 설정 편집에서 구성 입력을 선택한 채로 다음 구성을 입력합니다.

```
classification=spark-defaults,properties=[spark.executor.memory=2G]
```

5. 다른 옵션을 선택하고 항목을 선택한 다음 클러스터 생성을 선택합니다.

maximizeResourceAllocation을 설정하려면

- Spark가 설치된 클러스터를 생성하고 Amazon S3에 `myConfig.json`저장된 파일 AWS CLI를 참조하여를 사용하여 true로 `maximizeResourceAllocation` 설정합니다.

```
aws emr create-cluster --release-label emr-7.9.0 --applications Name=Spark \
--instance-type m5.xlarge --instance-count 2 --service-role EMR_DefaultRole_V2
 --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole --configurations https://
s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

`myConfig.json`:

```
[
  {
    "Classification": "spark",
    "Properties": {
      "maximizeResourceAllocation": "true"
    }
  }
]
```

> **Note**
>
> Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

Apache Log4j 1.x에서 Log4j 2.x로 마이그레이션

Apache Spark 릴리스 3.2.x 이하에서는 레거시 Apache Log4j 1.x 및 `log4j.properties` 파일을 사용하여 Spark 프로세스에서 Log4j를 구성합니다. Apache Spark 릴리스 3.3.0 이상에서는 Apache Log4j 2.x 및 `log4j2.properties` 파일을 사용하여 Spark 프로세스에서 Log4j를 구성합니다.

6.8.0 미만의 Amazon EMR 릴리스를 사용하여 Apache Spark Log4j를 구성한 경우 Amazon EMR 6.8.0 이상으로 업그레이드하기 전에 먼저 레거시 `spark-log4j` 구성 분류를 제거하고 `spark-log4j2` 구성 분류 및 키 형식으로 마이그레이션해야 합니다. 기존 `spark-log4j` 분류로 인해 Amazon EMR 릴리스 6.8.0 이상에서 `ValidationException` 오류로 클러스터 생성이 실패합니다. Log4j 비호환성과 관련된 오류에 대해서는 요금이 부과되지 않지만 계속하려면 더 이상 사용되지 않는 `spark-log4j` 구성 분류를 제거해야 합니다.

Apache Log4j 1.x에서 Log4j 2.x로 마이그레이션하는 방법에 대한 자세한 내용은 Github에서 Apache Log4j Migration Guide 및 Spark Log4j 2 Template을 참조하세요.

> **Note**
>
> Amazon EMR에서 Apache Spark는 Apache Log4j Migration Guide에 설명된 .xml 파일 대신 `log4j2.properties` 파일을 사용합니다. 또한 Log4j 1.x 브리지 메서드를 사용하여 Log4j 2.x로 변환하는 방법도 권장하지 않습니다.

Spark 성능 최적화

Amazon EMR은 Spark를 위한 여러 성능 최적화 성능을 제공합니다. 이 주제에서는 각 최적화 기능에 대해 자세히 설명합니다.

Spark 구성을 설정하는 방법에 대한 자세한 내용은 Spark 구성 단원을 참조하십시오.

적응형 쿼리 실행

적응형 쿼리 실행은 런타임 통계를 기반으로 쿼리 계획을 재최적화하기 위한 프레임워크입니다. Amazon EMR 5.30.0부터 Spark 2용 Apache Amazon EMR 런타임에서 Apache Spark 3의 다음과 같은 적응형 쿼리 실행 최적화를 사용할 수 있습니다.

- 적응형 조인 변환
- 셔플 파티션의 적응형 병합

적응형 조인 변환

적응형 조인 변환은 쿼리 스테이지의 런타임 크기에 따라 정렬-병합-조인 작업을 브로드캐스트-해시-조인 작업으로 변환하여 쿼리 성능을 개선합니다. 브로드캐스트-해시-조인은 조인의 한쪽이 출력을 모든 실행기에게 효율적으로 브로드캐스트할 수 있을 만큼 작은 경우 교환을 셔플하고 조인 양쪽을 모두

정렬할 필요가 없으므로 더 나은 성능을 지원합니다. 적응형 조인 변환은 Spark가 브로드캐스트-해시 조인을 자동으로 수행하는 경우더 다양한 사례를 지원합니다.

이 기능은 기본적으로 활성화되어 있습니다. `spark.sql.adaptive.enabled`를 `false`로 설정하여 비활성화할 수 있으며, 이렇게 하면 적응형 쿼리 실행 프레임워크도 비활성화됩니다. Spark는 조인 측 중 하나의 런타임 크기 통계가 `spark.sql.autoBroadcastJoinThreshold`(기본값: 10,485,760바이트(10MiB))를 초과하지 않을 때 정렬-병합-조인을 브로드캐스트-해시-조인으로 변환하기로 결정합니다.

셔플 파티션의 적응형 병합

셔플 파티션의 적응형 병합은 작은 연속 셔플 파티션을 통합하여 작은 작업이 너무 많아서 발생하는 오버헤드를 방지함으로써 쿼리 성능을 개선합니다. 이렇게 하면 더 많은 수의 초기 셔플 파티션을 미리 구성한 다음, 런타임 시 목표 크기로 줄일 수 있으므로 셔플 파티션을 더 균등하게 분산할 가능성이 커집니다.

이 기능은 명시적으로 `spark.sql.shuffle.partitions`를 설정하지 않는 한 기본적으로 활성화되어 있습니다. `spark.sql.adaptive.coalescePartitions.enabled`를 `true`로 설정하여 활성화할 수 있습니다. 셔플 파티션의 초기 수와 목표 파티션 크기는 각각 `spark.sql.adaptive.coalescePartitions.minPartitionNum` 및 `spark.sql.adaptive.advisoryPartitionSizeInBytes` 속성을 사용하여 조정할 수 있습니다. 이 기능의 관련 Spark 속성에 대한 자세한 내용은 다음 테이블을 참조하세요.

Spark 적응형 통합 파티션 속성

속성	기본값	설명
`spark.sql.adaptive.coalescePartitions.enabled`	true(명시적으로 `spark.sql.shuffle.partitions` 를 설정하지 않은 경우)	이 속성이 true이고 spark.sql.adaptive.enabled가 true인 경우 Spark는 목표 크기(`spark.sql.adaptive.advisoryPartitionSizeInBytes` 에서 지정함)에 따라 작은 작업이 너무 많이 발생하지 않도록 연속 셔플 파티션을 병합합니다.
`spark.sql.adaptive.advisoryPartitionSizeInBytes`	64MB	병합 시 셔플 파티션의 권장 크기(바이트). 이 구성은 `spark.sql.adaptive`

속성	기본값	설명
		.enabled 및 spark.sql.adaptive.coalescePartitions.enabled 모두 true인 경우에만 유효합니다.
spark.sql.adaptive.coalescePartitions.minPartitionNum	25	병합 후 셔플 파티션의 최소 수. 이 구성은 spark.sql.adaptive.enabled 및 spark.sql.adaptive.coalescePartitions.enabled 모두 true인 경우에만 유효합니다.
spark.sql.adaptive.coalescePartitions.initialPartitionNum	1000	병합 전 셔플 파티션의 초기 수. 이 구성은 spark.sql.adaptive.enabled 및 spark.sql.adaptive.coalescePartitions.enabled 모두 true인 경우에만 유효합니다.

동적 파티션 잘라내기

동적 파티션 잘라내기는 특정 쿼리를 위해 읽고 처리해야 하는 테이블 내 특정 파티션을 더 정확하게 선택하여 작업 성능을 향상시킵니다. 읽고 처리한 데이터의 양을 줄임으로써 작업 실행 시 상당한 시간이 절약됩니다. Amazon EMR 5.26.0의 경우 이 기능은 기본적으로 활성화되어 있습니다. Amazon EMR 5.24.0 및 5.25.0의 경우 Spark 내에서 또는 클러스터 생성 시 Spark 속성인 `spark.sql.dynamicPartitionPruning.enabled`를 설정하여 이 기능을 활성화할 수 있습니다.

Spark 동적 파티션 정리 파티션 속성

속성	기본 값	설명
`spark.sql.dynamicPartitionPruning.enabled`	`true`	true인 경우 동적 파티션 정리를 활성화합니다.
`spark.sql.optimizer.dynamicPartitionPruning.enforceBroadcastReuse`	`true`	`true`인 경우 Spark는 쿼리 실행 전에 방어 검사를 수행하여 동적 정리 필터에서의 브로드캐스트 교환 재사용이 사용자 정의 열 기반 규칙과 같은 이후 준비 규칙에 의해 중단되지 않도록 합니다. 재사용이 중단되고 이 구성이 `true`인 경우 Spark는 영향을 받는 동적 정리 필터를 제거하여 성능을 보호하고 정확성 문제를 방지합니다. 동적 정리 필터의 브로드캐스트 교환 시 해당 조인 작업의 브로드캐스트 교환과 다른, 일관되지 않은 결과가 생성되는 경우 정확성 문제가 발생할 수 있습니다. 이 구성을 `false`로 설정할 때는 주의해야 합니다. 이렇게 하면 사용자 정의 열 기반 규칙에 의해 재사용이 중단되는 경우와 같은 시나리오를 해결할 수 있습니다. 적응형 쿼리 실행이 활성화되면 브로드캐스트 재사용이 항상 적용됩니다.

이러한 최적화를 통해 계획 시간에 해결할 수 있는 정적 조건자를 푸시다운하는 것만 지원하는 Spark 2.4.2의 기존 기능보다 나은 결과를 얻을 수 있습니다.

다음은 Spark 2.4.2이 실행하는 정적 조건자 푸시다운의 예입니다.

```
partition_col = 5

partition_col IN (1,3,5)

partition_col between 1 and 3

partition_col = 1 + 3
```

동적 파티션 잘라내기를 통해 Spark 엔진은 읽어야 할 파티션과 안전하게 제거할 수 있는 파티션을 실행 시간에 동적으로 추론할 수 있습니다. 예를 들어 다음 쿼리에는 두 가지 테이블이 수반됩니다. 즉 모든 스토어에 대해 총 매출 전체를 포함하고 리전에 따라 분할된 `store_sales` 테이블과 각 국가에 대해 리전 매핑을 포함하는 `store_regions` 테이블입니다. 테이블에는 전 세계에 분산된 스토어에 대한 데이터가 포함되어 있지만, 여기서는 북미에 대한 데이터만 쿼리합니다.

```
select ss.quarter, ss.region, ss.store, ss.total_sales
from store_sales ss, store_regions sr
where ss.region = sr.region and sr.country = 'North America'
```

동적 파티션 잘라내기가 없으면 이 쿼리는 하위 쿼리의 결과와 일치하는 리전의 하위 집합을 필터링하기 전에 모든 리전을 읽습니다. 동적 파티션 잘라내기가 있는 경우 이 쿼리는 하위 쿼리에서 반환되는 리전에 대한 파티션만 읽고 처리합니다. 이를 통해 스토리지에서 더 적은 데이터를 읽고 더 적은 레코드를 처리함으로써 시간과 리소스를 절약합니다.

스칼라 하위 쿼리 평면화

이 최적화를 통해 동일한 테이블에 대한 스칼라 하위 쿼리가 있는 쿼리의 성능이 향상됩니다. Amazon EMR 5.26.0의 경우 이 기능은 기본적으로 활성화되어 있습니다. Amazon EMR 5.24.0 및 5.25.0의 경우 Spark 내에서 또는 클러스터 생성 시 Spark 속성인 `spark.sql.optimizer.flattenScalarSubqueriesWithAggregates.enabled`를 설정하여 이 기능을 활성화할 수 있습니다. 이 속성을 true로 설정하면 이 쿼리 최적화 프로그램은 가능한 경우 동일한 관계를 사용하는 총 스칼라 하위 쿼리를 평면화합니다. 스칼라 하위 쿼리는 집계 함수에 대한 하위 쿼리에 있는 모든 조건자를 푸시한 후 관계당 모든 집계 함수로 한 가지 집계를 수행하는 방식으로 평면화됩니다.

다음은 이러한 최적화를 통해 이익을 얻을 쿼리의 샘플입니다.

```
select (select avg(age) from students                    /* Subquery 1 */
```

```
                   where age between 5 and 10) as group1,
        (select avg(age) from students                    /* Subquery 2 */
                   where age between 10 and 15) as group2,
        (select avg(age) from students                    /* Subquery 3 */
                   where age between 15 and 20) as group3
```

최적화를 통해 이전 쿼리는 다음과 같이 재작성됩니다.

```
select c1 as group1, c2 as group2, c3 as group3
from (select avg (if(age between 5 and 10, age, null)) as c1,
             avg (if(age between 10 and 15, age, null)) as c2,
             avg (if(age between 15 and 20, age, null)) as c3 from students);
```

재작성된 쿼리는 학생 테이블을 한 번만 읽고 세 가지 하위 쿼리의 조건자는 avg 함수로 푸시된다는 점에 유의하십시오.

INTERSECT에 앞선 DISTINCT

이 최적화에서는 INTERSECT 사용 시 조인을 최적화합니다. Amazon EMR 5.26.0의 경우 이 기능은 기본적으로 활성화되어 있습니다. Amazon EMR 5.24.0 및 5.25.0의 경우 Spark 내에서 또는 클러스터 생성 시 Spark 속성인 `spark.sql.optimizer.distinctBeforeIntersect.enabled`를 설정하여 이 기능을 활성화할 수 있습니다. INTERSECT를 사용하는 쿼리는 자동으로 좌측 세미 조인을 사용하도록 자동 변환됩니다. 이 속성이 true로 설정된 경우 DISTINCT 연산자가 좌측 세미 조인을 SortMergeJoin이 아닌 BroadcastHashJoin으로 만들 수 있음을 감지한 쿼리 최적화 프로그램은 DISTINCT 연산자를 INTERSECT의 하위 항목으로 푸시합니다.

다음은 이러한 최적화를 통해 이익을 얻을 쿼리의 샘플입니다.

```
(select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
intersect
(select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
```

이 `spark.sql.optimizer.distinctBeforeIntersect.enabled` 속성을 활성화하지 않으면 쿼리는 다음과 같이 재작성됩니다.

```
select distinct brand from
  (select item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
```

```
left semi join
   (select item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
 on brand <=> cs_brand
```

이 spark.sql.optimizer.distinctBeforeIntersect.enabled 속성을 활성화하면 쿼리는 다음과 같이 재작성됩니다.

```
select brand from
  (select distinct item.brand brand from store_sales, item
     where store_sales.item_id = item.item_id)
left semi join
   (select distinct item.brand cs_brand from catalog_sales, item
     where catalog_sales.item_id = item.item_id)
 on brand <=> cs_brand
```

블룸 필터 조인

이러한 최적화는 조인의 다른 측면 값에서 생성된 블룸 필터를 사용해 조인의 한 측면을 사전 필터링하여 일부 조인의 성능을 향상시킬 수 있습니다. Amazon EMR 5.26.0의 경우 이 기능은 기본적으로 활성화되어 있습니다. Amazon EMR 5.25.0의 경우 Spark 내에서 또는 클러스터 생성 시 Spark 속성 spark.sql.bloomFilterJoin.enabled를 true로 설정하여 이 기능을 활성화할 수 있습니다.

다음은 블룸 필터를 활용할 수 있는 예제 쿼리입니다.

```
select count(*)
from sales, item
where sales.item_id = item.id
and item.category in (1, 10, 16)
```

이 기능을 사용할 때, 블룸 필터는 카테고리가 쿼리 처리되는 카테고리 세트에 있는 모든 항목 id들로 구축됩니다. 매출 테이블을 스캔하는 동안 블룸 필터를 사용해 어떤 매출이 블룸 필터에서 지정한 세트에서 분명 제외된 항목에 있는지 파악합니다. 따라서 이렇게 확인된 매출은 가능한 한 조기에 필터링될 수 있습니다.

최적화된 조인 재정렬

이 최적화로 필터와 함께 테이블에 관여하는 조인들을 재정렬하여 쿼리 성능을 향상시킬 수 있습니다. Amazon EMR 5.26.0의 경우 이 기능은 기본적으로 활성화되어 있습니다. Amazon EMR 5.25.0의 경우 Spark 구성 파라미터 spark.sql.optimizer.sizeBasedJoinReorder.enabled를 true로 설정

하여 이 기능을 활성화할 수 있습니다. Spark의 기본 동작은 쿼리에 기재된 대로 테이블을 왼쪽에서 오른쪽으로 조인하는 것입니다. 이러한 전략은 많이 사용하는 조인을 나중에 활용하기 위해 필터가 있는 소규모 조인을 우선 실행할 수 있는 기회를 놓칠 수 있습니다.

아래의 예시 쿼리는 한 국가의 모든 매장에 있는 모든 반환품을 보고합니다. 조인 재정렬을 최적화하지 않으면, Spark는 두 개의 대형 테이블인 `store_sales`와 `store_returns`를 우선 조인한 다음, 두 대형 테이블을 `store` 그리고 마지막으로 `item`과 조인합니다.

```
select ss.item_value, sr.return_date, s.name, i.desc,
from store_sales ss, store_returns sr, store s, item i
where ss.id = sr.id and ss.store_id = s.id and ss.item_id = i.id
and s.country = 'USA'
```

최적화된 조인 재정렬을 활용해 Spark는 `store_sales`와 `store`를 우선 조인하는 데, 이는 `store`에 필터가 있으며 `store_returns` 및 `broadcastable`보다 소규모이기 때문입니다. 그런 다음 Spark는 `store_returns`와 조인한 뒤 마지막으로 `item`와 조인합니다. `item` 에 필터가 있고 브로드캐스팅할 수 있다면, 재정렬에 적격하며 `store_sales`를 `store`와, 그리고 `item`, 또 결국 `store_returns`와 조인하게 됩니다.

Spark 결과 조각 캐싱

Amazon EMR 6.6.0 이상에는 결과 조각을 자동으로 캐싱하는 선택적 Spark 결과 조각 캐싱 기능이 포함되어 있습니다. 이러한 결과 조각은 선택한 Amazon S3 버킷에 저장된 쿼리 하위 트리에서 얻은 결과의 일부입니다. 저장된 쿼리 결과 조각은 후속 쿼리 실행 시 재사용되므로 쿼리 속도가 빨라집니다.

결과 조각 캐싱은 Spark 쿼리를 분석하고 지정된 S3 위치에 적합한 결과 조각을 캐싱합니다. 후속 쿼리 실행 시 사용 가능한 쿼리 결과 조각을 자동으로 감지하여 S3에서 가져옵니다. 결과 조각 캐싱은 후속 쿼리가 원래 쿼리와 정확히 일치해야 캐시에서 결과를 반환하는 결과 세트 캐싱과 다릅니다. 데이터의 정적 하위 세트를 반복적으로 대상으로 하는 쿼리에 사용할 경우 결과 조각 캐싱은 성능을 크게 개선합니다.

2022년까지 주문을 집계하는 다음 쿼리를 고려합니다.

```
select
    l_returnflag,
    l_linestatus,
    count(*) as count_order
from
    lineitem
where
```

```
    l_shipdate <= current_date
    and year(l_shipdate) == '2022'
group by
    l_returnflag,
    l_linestatus
```

시간이 지남에 따라 이 쿼리를 매일 실행하여 해당 연도의 총 판매를 보고해야 합니다. 결과 조각 캐싱을 사용하지 않으면 해당 연도의 모든 날짜에 대한 결과를 매일 다시 계산해야 합니다. 쿼리 속도는 시간이 지날수록 느려지고 365일 분량의 결과를 모두 다시 계산해야 하는 연말에는 속도가 가장 느려집니다.

결과 조각 캐싱을 활성화하면 캐시에 있는 해당 연도의 모든 이전 날짜에 대한 결과를 사용합니다. 이 기능은 매일 하루의 결과만 다시 계산하면 됩니다. 기능이 결과 조각을 계산한 후 해당 조각을 캐싱합니다. 따라서 캐시를 사용하는 쿼리 시간이 빨라지고 후속 쿼리마다 일정한 시간이 유지됩니다.

Spark 결과 조각 캐싱 활성화

Spark 결과 조각 캐싱을 활성화하려면 다음 단계를 수행합니다.

1. Amazon S3에 캐시 버킷을 생성하고 EMRFS에 대한 읽기 및 쓰기 액세스를 승인합니다. 자세한 내용은 Amazon S3에서 EMRFS 데이터에 대한 액세스 권한 부여 단원을 참조하십시오.
2. 기능을 활성화하려면 Amazon EMR Spark 구성을 설정합니다.

```
spark.subResultCache.enabled = true
spark.subResultCache.fs.root.path = s3://&example-s3-bucket;/cache_dir/
```

3. 버킷의 S3 수명 주기 관리를 활성화하여 캐시 파일을 자동으로 정리합니다.
4. 선택적으로 reductionRationThreshold 및 maxBufferSize 속성을 구성하여 기능을 추가로 조정합니다.

```
spark.sql.subResultCache.reductionRatioThreshold
spark.sql.subResultCache.maxBufferSize
```

결과 조각 캐싱 사용 시 고려 사항

다시 계산하는 대신 Amazon S3에 이미 캐시된 결과를 사용할 경우 동일한 캐시된 결과를 사용할 수 있는 횟수만큼 비용 절감 효과가 커집니다. 대형 테이블을 스캔한 후 결과 크기를 8배 이상 줄이는 필터 또는 해시 집계(즉, 입력 크기 대 결과의 비율이 8:1 이상인 경우)를 사용하는 쿼리에서는 이 기능을 최

대한 활용할 수 있습니다. 입력과 결과 간의 감소 비율이 클수록 비용상의 이점도 커집니다. 감소 비율은 작지만 테이블 스캔과 필터 또는 집계 사이에 비용이 많이 드는 계산 단계가 포함된 쿼리도 결과를 생성하는 데 드는 비용이 Amazon S3에서 결과를 가져오는 데 드는 비용보다 큰 경우 비용이 절감됩니다. 기본적으로 결과 조각 캐싱은 감소 비율이 8:1 이상일 것으로 감지되는 경우에만 유효합니다.

쿼리가 캐시된 결과를 반복적으로 재사용하는 경우 이 기능의 이점이 가장 큽니다. 롤링 및 증분 기간 쿼리가 좋은 예입니다. 예를 들어 이미 29일 동안 실행된 30일의 롤링 기간 쿼리는 원래 입력 소스에서 대상 데이터의 1/30만 가져오면 되고 이전 29일 동안 캐시된 결과 조각을 사용합니다. 증분 기간 쿼리는 기간의 시작 위치가 고정되어 있기 때문에 훨씬 더 유용합니다. 쿼리를 간접 호출할 때마다 입력 소스에서 읽어야 하는 처리 비율이 줄어듭니다.

결과 조각 캐싱을 사용할 때 고려해야 할 추가 사항은 다음과 같습니다.

- 동일한 쿼리 조각을 포함하는 동일한 데이터를 대상으로 하지 않는 쿼리는 캐시 적중률이 낮으므로 이 기능을 활용할 수 없습니다.
- 비용이 많이 드는 계산 단계를 포함하지 않고 감소 비율이 낮은 쿼리에서는 캐시된 결과에서는 처음에 처리할 때와 읽기 비용이 거의 같습니다.
- 캐시에 쓰는 데 드는 비용 때문에 첫 번째 쿼리에서는 항상 약간의 회귀가 나타납니다.
- 결과 조각 캐싱 기능은 Parquet 파일에서만 작동합니다. 다른 파일 형식은 지원되지 않습니다.
- 결과 조각 캐싱 기능 버퍼는 파일 분할 크기가 128MB 이상인 스캔만 캐시하려고 시도합니다. 기본 Spark 구성에서 스캔 크기(스캔 중인 모든 파일의 총 크기)를 실행기 코어 수로 나눈 값이 128MB 미만이면 결과 조각 캐싱이 비활성화됩니다. 아래 나열된 Spark 구성 중 하나를 설정하는 경우 파일 분할 크기는 다음과 같습니다.

  ```
  min(maxPartitionBytes, max(openCostInBytes, scan size / minPartitionNum))
  ```

  - spark.sql.leafNodeDefaultParallelism(기본값: spark.default.parallelism)
  - spark.sql.files.minPartitionNum(기본값: spark.sql.leafNodeDefaultParallelism)
  - spark.sql.files.openCostInBytes
  - spark.sql.files.maxPartitionBytes
- 결과 조각 캐싱 기능은 RDD 파티션 단위로 캐싱합니다. 앞에서 설명한 감소 비율(기본값 8:1)이 RDD 파티션별로 평가됩니다. RDD당 감소 비율이 8:1보다 크거나 작은 워크로드는 RDD당 감소 비율이 지속적으로 8:1 미만인 워크로드보다 성능상의 이점이 적을 수 있습니다.
- 결과 조각 캐싱 기능은 캐시되는 각 RDD 파티션에 대해 기본적으로 16MB 쓰기 버퍼를 사용합니다. RDD 파티션당 16MB가 넘게 캐시될 경우 쓰기 불가능 여부를 판단하는 데 드는 비용으로 인해 성능이 저하될 수 있습니다.

- 기본적으로 결과 조각 캐싱은 감소 비율이 8:1 미만인 RDD 파티션 결과를 캐시하려고 시도하지 않으며, 쓰기 버퍼를 16MB로 제한하지만 다음 구성을 통해 이 두 값을 모두 조정할 수 있습니다.

```
spark.sql.subResultCache.reductionRatioThreshold (default: 8.0)
spark.sql.subResultCache.maxBufferSize (default: 16MB, max: 64MB)
```

- 동일한 Amazon EMR 릴리스를 사용하는 여러 클러스터는 동일한 캐시 위치를 공유할 수 있습니다. 결과의 정확성을 보장하기 위해 결과 조각 캐싱은 여러 Amazon EMR 릴리스에서 작성된 캐시 결과를 사용하지 않습니다.
- 결과 조각 캐싱은 Spark 스트리밍 사용 사례 또는 RecordServer, Apache Ranger 또는가 사용되는 경우 자동으로 비활성화 AWS Lake Formation 됩니다.
- 결과 조각 캐시 읽기/쓰기는 EMRFS/S3A 및 Amazon S3 버킷을 사용합니다. CSE(EMRFS만 해당)/SSE S3/ SSE KMS 암호화가 지원됩니다. 컨텍스트를 위해 S3A는 클러스터가 Amazon S3에서 데이터를 읽고 쓸 수 있도록 하둡 구현을 제공합니다. S3A에 대한 지원은 EMR-7.4.0 이상에서 사용할 수 있습니다.

Apache Spark용 Nvidia RAPIDS 액셀러레이터 사용

Amazon EMR 릴리스 6.2.0 이상에서는 Nvidia의 Apache Spark용 RAPIDS 액셀러레이터 플러그인을 통해 EC2 그래픽 처리 장치(GPU) 인스턴스 유형을 사용하여 Spark를 가속화할 수 있습니다. RAPIDS Accelerator는 코드 변경 없이 Apache Spark 3.0 데이터 과학 파이프라인에서 GPU를 가속화하고, 데이터 처리 및 모델 교육 속도를 높이는 동시에 인프라 비용을 크게 절감합니다.

다음 섹션에서는 Spark용 Spark RAPIDS 플러그인을 사용하도록 EMR 클러스터를 구성하는 방법을 안내합니다.

인스턴스 유형 선택

Spark용 Nvidia Spark RAPIDS 플러그인을 사용하려면 코어 및 태스크 인스턴스 그룹에서 Spark-RAPIDS의 하드웨어 요구 사항을 충족하는 EC2 GPU 인스턴스 유형을 사용해야 합니다. Amazon EMR에서 지원되는 GPU 인스턴스 유형의 전체 목록을 보려면 Amazon EMR 관리 안내서에서 지원되는 인스턴스 유형을 참조하세요. 기본 인스턴스 그룹의 인스턴스 유형은 GPU 또는 비GPU 유형일 수 있지만 ARM 인스턴스 유형은 지원되지 않습니다.

클러스터의 애플리케이션 구성 설정

1: Amazon EMR을 활성화하여 새 클러스터에 플러그인 설치

플러그인을 설치하려면 클러스터를 생성할 때 다음 구성을 제공합니다.

```
{
 "Classification":"spark",
 "Properties":{
  "enableSparkRapids":"true"
 }
}
```

2. GPU를 사용하도록 YARN 구성

YARN 기반 GPU를 사용하는 방법에 대한 자세한 내용은 Apache Hadoop 설명서에서 [YARN에서 GPU 사용]을 참조하세요. 다음 예에서는 Amazon EMR 6.x 및 7.x 릴리스를 위한 샘플 YARN 구성을 보여줍니다.

Amazon EMR 7.x

Amazon EMR 7.x용 YARN 구성 예제

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
```

```
                "module.enabled":"true"
            }
        },
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/spark-rapids-cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
}
```

Amazon EMR 6.x

Amazon EMR 6.x용 YARN 구성 예제

```
{
    "Classification":"yarn-site",
    "Properties":{
        "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
        "yarn.resource-types":"yarn.io/gpu",
        "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
        "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/usr/
bin",
        "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
        "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
        "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
        "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
    }
},{
    "Classification":"container-executor",
    "Properties":{

    },
    "Configurations":[
        {
            "Classification":"gpu",
            "Properties":{
                "module.enabled":"true"
            }
        },
```

```
        {
            "Classification":"cgroups",
            "Properties":{
                "root":"/sys/fs/cgroup",
                "yarn-hierarchy":"yarn"
            }
        }
    ]
}
```

3. RAPIDS를 사용하도록 Spark 구성

Spark에서 RAPIDS 플러그인을 사용하려면 다음 구성이 필요합니다.

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.plugins":"com.nvidia.spark.SQLPlugin",
  "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/gpu/
getGpusResources.sh",
  "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-linux/lib:/usr/
local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/cuda/lib:/usr/
local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/native:/docker/usr/
lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native"
 }
}
```

XGBoost 설명서의 XGBoost4J-Spark 라이브러리는 클러스터에서 Spark RAPIDS 플러그인이 활성화된 경우에도 사용할 수 있습니다. 다음 구성을 사용하여 XGBoost를 Spark 작업과 통합할 수 있습니다.

```
{
 "Classification":"spark-defaults",
 "Properties":{
  "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-spark_3.0-1.4.2-0.3.0.jar"
 }
}
```

GPU 가속 EMR 클러스터를 튜닝하는 데 사용할 수 있는 추가 Spark 구성은 Nvidia.github.io 설명서에서 Rapids Accelerator for Apache Spark 튜닝 안내서를 참조하세요.

4. YARN 용량 스케줄러 구성

GPU 예약 및 격리를 활성화하도록 `DominantResourceCalculator`를 구성해야 합니다. 자세한 내용은 Apache Hadoop 설명서에서 Using GPU on YARN을 참조하세요.

```
{
 "Classification":"capacity-scheduler",
 "Properties":{
  "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
 }
}
```

5. 구성을 포함하도록 JSON 파일 생성

Spark 클러스터용 RAPIDS 플러그인을 사용하기 위한 구성이 포함된 JSON 파일을 생성할 수 있습니다. 나중에 클러스터를 시작할 때 파일을 제공합니다.

파일을 로컬 또는 S3에 저장할 수 있습니다. 클러스터에 애플리케이션 구성을 제공하는 방법에 대한 자세한 내용은 애플리케이션 구성 섹션을 참조하세요.

다음 샘플 파일을 템플릿으로 사용하고 고유한 구성을 만드세요.

Amazon EMR 7.x

Amazon EMR 7.x용 예제 **`my-configurations.json`** 파일

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/spark-
rapids-cgroup",
```

```
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
    {
        "Classification":"container-executor",
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/spark-rapids-cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
            "spark.task.cpus":"1",
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
```

```
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
    },
    {
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
]
```

Amazon EMR 6.x

Amazon EMR 6.x용 예제 **my-configurations.json** 파일

```
[
    {
        "Classification":"spark",
        "Properties":{
            "enableSparkRapids":"true"
        }
    },
    {
        "Classification":"yarn-site",
        "Properties":{
            "yarn.nodemanager.resource-plugins":"yarn.io/gpu",
            "yarn.resource-types":"yarn.io/gpu",
            "yarn.nodemanager.resource-plugins.gpu.allowed-gpu-devices":"auto",
            "yarn.nodemanager.resource-plugins.gpu.path-to-discovery-executables":"/
usr/bin",
            "yarn.nodemanager.linux-container-executor.cgroups.mount":"true",
            "yarn.nodemanager.linux-container-executor.cgroups.mount-path":"/sys/fs/
cgroup",
            "yarn.nodemanager.linux-container-executor.cgroups.hierarchy":"yarn",
            "yarn.nodemanager.container-
executor.class":"org.apache.hadoop.yarn.server.nodemanager.LinuxContainerExecutor"
        }
    },
    {
        "Classification":"container-executor",
```

```
        "Properties":{

        },
        "Configurations":[
            {
                "Classification":"gpu",
                "Properties":{
                    "module.enabled":"true"
                }
            },
            {
                "Classification":"cgroups",
                "Properties":{
                    "root":"/sys/fs/cgroup",
                    "yarn-hierarchy":"yarn"
                }
            }
        ]
    },
    {
        "Classification":"spark-defaults",
        "Properties":{
            "spark.plugins":"com.nvidia.spark.SQLPlugin",
            "spark.executor.resource.gpu.discoveryScript":"/usr/lib/spark/scripts/
gpu/getGpusResources.sh",
            "spark.executor.extraLibraryPath":"/usr/local/cuda/targets/x86_64-
linux/lib:/usr/local/cuda/extras/CUPTI/lib64:/usr/local/cuda/compat/lib:/usr/local/
cuda/lib:/usr/local/cuda/lib64:/usr/lib/hadoop/lib/native:/usr/lib/hadoop-lzo/lib/
native:/docker/usr/lib/hadoop/lib/native:/docker/usr/lib/hadoop-lzo/lib/native",
            "spark.submit.pyFiles":"/usr/lib/spark/jars/xgboost4j-
spark_3.0-1.4.2-0.3.0.jar",
            "spark.rapids.sql.concurrentGpuTasks":"1",
            "spark.executor.resource.gpu.amount":"1",
            "spark.executor.cores":"2",
            "spark.task.cpus":"1",
            "spark.task.resource.gpu.amount":"0.5",
            "spark.rapids.memory.pinnedPool.size":"0",
            "spark.executor.memoryOverhead":"2G",
            "spark.locality.wait":"0s",
            "spark.sql.shuffle.partitions":"200",
            "spark.sql.files.maxPartitionBytes":"512m"
        }
    },
    {
```

```
        "Classification":"capacity-scheduler",
        "Properties":{
            "yarn.scheduler.capacity.resource-
calculator":"org.apache.hadoop.yarn.util.resource.DominantResourceCalculator"
        }
    }
]
```

클러스터에 대한 부트스트랩 작업 추가

클러스터를 생성할 때 부트스트랩 작업 스크립트를 제공하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 부트스트랩 작업 기본 사항을 참조하세요.

다음 예제 스크립트에서는 Amazon EMR 6.x 및 7.x용 부트스트랩 작업 파일을 만드는 방법을 확인할 수 있습니다.

Amazon EMR 7.x

Amazon EMR 7.x용 예제 **my-bootstrap-action.sh** 파일

YARN을 사용하여 Amazon EMR 7.x 릴리스에서 GPU 리소스를 관리하기 위해서는 클러스터에 CGroup v1을 수동으로 마운트해야 합니다. 이 예제처럼 부트스트랩 작업 스크립트를 이용하여 이 작업을 할 수 있습니다.

```
#!/bin/bash
set -ex

sudo mkdir -p /spark-rapids-cgroup/devices
sudo mount -t cgroup -o devices cgroupv1-devices /spark-rapids-cgroup/devices
sudo chmod a+rwx -R /spark-rapids-cgroup
```

Amazon EMR 6.x

Amazon EMR 6.x용 예제 **my-bootstrap-action.sh** 파일

Amazon EMR 6.x 릴리스의 경우에는 클러스터에서 YARN에 대한 CGroup 권한을 열어야 합니다. 이 예제처럼 부트스트랩 작업 스크립트를 이용하여 이 작업을 할 수 있습니다.

```
#!/bin/bash
set -ex
```

```
sudo chmod a+rwx -R /sys/fs/cgroup/cpu,cpuacct
sudo chmod a+rwx -R /sys/fs/cgroup/devices
```

클러스터 시작

마지막 단계는 위에서 언급한 클러스터 구성으로 클러스터를 시작하는 것입니다. 다음은 Amazon EMR CLI에서 클러스터를 시작하는 예제 명령입니다.

```
 aws emr create-cluster \
--release-label emr-7.9.0 \
--applications Name=Hadoop Name=Spark \
--service-role EMR_DefaultRole_V2 \
--ec2-attributes KeyName=my-key-pair,InstanceProfile=EMR_EC2_DefaultRole \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m4.4xlarge \

                  InstanceGroupType=CORE,InstanceCount=1,InstanceType=g4dn.2xlarge \

                  InstanceGroupType=TASK,InstanceCount=1,InstanceType=g4dn.2xlarge \
--configurations file:///my-configurations.json \
--bootstrap-actions Name='My Spark Rapids Bootstrap action',Path=s3://amzn-s3-demo-
bucket/my-bootstrap-action.sh
```

Spark 쉘에 액세스

Spark 쉘은 Scala REPL (Read-Eval-Print-Loop)을 기반으로 합니다. Scala REPL을 사용하면 Spark 프로그램을 대화식으로 생성하고 작업을 프레임워크로 제출할 수 있습니다. 프라이머리 노드를 SSH에 연결하고 `spark-shell`을 간접 호출하여 Spark 쉘에 액세스할 수 있습니다. SSH를 사용하여 프라이머리 노드에 연결하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요. 다음 예제에서는 Amazon S3에 저장된 Apache HTTP Server 액세스 로그가 사용됩니다.

> **Note**
>
> 이러한 예제의 Amazon 버킷은 미국 동부(버지니아 북부)에 액세스할 수 있는 클라이언트에서 사용 가능합니다.

기본적으로 Spark 쉘은 `sc`라는 자체 SparkContext 객체를 생성합니다. REPL 내에서 필요한 경우 이 컨텍스트를 사용할 수 있습니다. 또한 sqlContext도 이 쉘에서 사용할 수 있으며, HiveContext입니다.

Example Spark 쉘을 사용하여 Amazon S3에 저장된 파일에서 특정 문자열의 발생 개수 계산

이 예제에서는 sc를 사용하여 Amazon S3에 저장된 텍스트 파일을 읽습니다.

```
scala> sc
res0: org.apache.spark.SparkContext = org.apache.spark.SparkContext@404721db

scala> val textFile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/
impressions/dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark에서 textFile 및 관련 [데이터 구조]가 생성됩니다. 그 다음, 이 예제는 로그 파일에서 "cartoonnetwork.com" 문자열이 나오는 행 수를 셉니다.

```
scala> val linesWithCartoonNetwork = textFile.filter(line =>
 line.contains("cartoonnetwork.com")).count()
linesWithCartoonNetwork: org.apache.spark.rdd.RDD[String] = MapPartitionsRDD[2] at
 filter at <console>:23
<snip>
<Spark program runs>
scala> linesWithCartoonNetwork
res2: Long = 9
```

Example Python 기반 Spark 쉘을 사용하여 Amazon S3에 저장된 파일에서 특정 문자열의 발생 개수 계산

Spark에는 Python 기반 셸인 pyspark도 포함되어 있으며, 이 셸을 사용하여 Python에서 작성된 Spark 프로그램을 시제품화할 수 있습니다. spark-shell과 마찬가지로, 프라이머리 노드에서 pyspark를 간접 호출합니다. 이 쉘에도 동일한 SparkContext 객체가 있습니다.

```
>>> sc
<pyspark.context.SparkContext object at 0x7fe7e659fa50>
>>> textfile = sc.textFile("s3://elasticmapreduce/samples/hive-ads/tables/impressions/
dt=2009-04-13-08-05/ec2-0-51-75-39.amazon.com-2009-04-13-08-05.log")
```

Spark에서 textFile 및 관련 데이터 구조가 생성됩니다. 그 다음, 이 예제는 로그 파일에서 "cartoonnetwork.com" 문자열이 나오는 행 수를 셉니다.

```
>>> linesWithCartoonNetwork = textfile.filter(lambda line: "cartoonnetwork.com" in
 line).count()
15/06/04 17:12:22 INFO lzo.GPLNativeCodeLoader: Loaded native gpl library from the
 embedded binaries
```

```
15/06/04 17:12:22 INFO lzo.LzoCodec: Successfully loaded & initialized native-lzo
 library [hadoop-lzo rev EXAMPLE]
15/06/04 17:12:23 INFO fs.EmrFileSystem: Consistency disabled, using
 com.amazon.ws.emr.hadoop.fs.s3n.S3NativeFileSystem as filesystem implementation
<snip>
<Spark program continues>
>>> linesWithCartoonNetwork
9
```

기계 학습에 Amazon SageMaker Spark 사용

Amazon EMR 릴리스 5.11.0 이상을 사용하면 Spark와 함께 `aws-sagemaker-spark-sdk` 구성 요소가 설치됩니다. 이 구성 요소는 Amazon SageMaker Spark 및 Spark와 Amazon SageMaker의 통합을 위한 관련 종속성을 설치합니다. Amazon SageMaker Spark를 사용하면 Amazon SageMaker 단계를 사용하여 Spark 기계 학습(ML) 파이프라인을 생성할 수 있습니다. 자세한 내용은 GitHub에서 Amazon SageMaker Spark README와 Amazon SageMaker 개발자 안내서에서 Using Apache Spark with Amazon SageMaker를 참조하세요.

Spark 애플리케이션 작성

Spark 애플리케이션은 Scala, Java, 또는 Python으로 작성할 수 있습니다. Apache Spark 설명서에서 Spark examples 주제에 여러 Spark 애플리케이션 예제가 있습니다. 아래에는 기본적으로 지원되는 세 가지 애플리케이션 중 Estimating Pi 예제가 표시되어 있습니다. 또한 `$SPARK_HOME/examples` 및 GitHub에서 전체 예제를 볼 수 있습니다. Spark용 JAR 빌드 방법에 대한 자세한 내용은 Apache Spark 문서에서 Quick Start 주제를 참조하세요.

Scala

Scala 호환성 문제를 방지하려면 Amazon EMR 클러스터용 Spark 애플리케이션을 컴파일할 때 올바른 Scala 버전에 대한 Spark 종속성을 사용하는 것이 좋습니다. 사용해야 하는 Scala 버전은 클러스터에 설치된 Spark 버전에 따라 다릅니다. 예를 들어, Amazon EMR 릴리스 5.30.1은 Scala 2.11과 함께 빌드된 Spark 2.4.5를 사용합니다. 클러스터에서 Amazon EMR 릴리스 5.30.1을 사용하는 경우 Scala 2.11의 Spark 종속성을 사용합니다. Spark에서 사용하는 Scala 버전에 대한 자세한 내용은 Apache Spark 설명서를 참조하세요.

```
package org.apache.spark.examples
import scala.math.random
import org.apache.spark._
```

```
/** Computes an approximation to pi */
object SparkPi {
  def main(args: Array[String]) {
    val conf = new SparkConf().setAppName("Spark Pi")
    val spark = new SparkContext(conf)
    val slices = if (args.length > 0) args(0).toInt else 2
    val n = math.min(100000L * slices, Int.MaxValue).toInt // avoid overflow
    val count = spark.parallelize(1 until n, slices).map { i =>
      val x = random * 2 - 1
      val y = random * 2 - 1
      if (x*x + y*y < 1) 1 else 0
    }.reduce(_ + _)
    println("Pi is roughly " + 4.0 * count / n)
    spark.stop()
  }
}
```

Java

```
package org.apache.spark.examples;

import org.apache.spark.SparkConf;
import org.apache.spark.api.java.JavaRDD;
import org.apache.spark.api.java.JavaSparkContext;
import org.apache.spark.api.java.function.Function;
import org.apache.spark.api.java.function.Function2;

import java.util.ArrayList;
import java.util.List;

/**
 * Computes an approximation to pi
 * Usage: JavaSparkPi [slices]
 */
public final class JavaSparkPi {

  public static void main(String[] args) throws Exception {
    SparkConf sparkConf = new SparkConf().setAppName("JavaSparkPi");
    JavaSparkContext jsc = new JavaSparkContext(sparkConf);

    int slices = (args.length == 1) ? Integer.parseInt(args[0]) : 2;
    int n = 100000 * slices;
    List<Integer> l = new ArrayList<Integer>(n);
```

```
    for (int i = 0; i < n; i++) {
      l.add(i);
    }

    JavaRDD<Integer> dataSet = jsc.parallelize(l, slices);

    int count = dataSet.map(new Function<Integer, Integer>() {
      @Override
      public Integer call(Integer integer) {
        double x = Math.random() * 2 - 1;
        double y = Math.random() * 2 - 1;
        return (x * x + y * y < 1) ? 1 : 0;
      }
    }).reduce(new Function2<Integer, Integer, Integer>() {
      @Override
      public Integer call(Integer integer, Integer integer2) {
        return integer + integer2;
      }
    });

    System.out.println("Pi is roughly " + 4.0 * count / n);

    jsc.stop();
  }
}
```

Python

```

import argparse
import logging
from operator import add
from random import random

from pyspark.sql import SparkSession

logger = logging.getLogger(__name__)
logging.basicConfig(level=logging.INFO, format="%(levelname)s: %(message)s")


def calculate_pi(partitions, output_uri):
    """
    Calculates pi by testing a large number of random numbers against a unit circle
```

```
    inscribed inside a square. The trials are partitioned so they can be run in
    parallel on cluster instances.

    :param partitions: The number of partitions to use for the calculation.
    :param output_uri: The URI where the output is written, typically an Amazon S3
                       bucket, such as 's3://example-bucket/pi-calc'.
    """

    def calculate_hit(_):
        x = random() * 2 - 1
        y = random() * 2 - 1
        return 1 if x**2 + y**2 < 1 else 0

    tries = 100000 * partitions
    logger.info(
        "Calculating pi with a total of %s tries in %s partitions.", tries, partitions
    )
    with SparkSession.builder.appName("My PyPi").getOrCreate() as spark:
        hits = (
            spark.sparkContext.parallelize(range(tries), partitions)
            .map(calculate_hit)
            .reduce(add)
        )
        pi = 4.0 * hits / tries
        logger.info("%s tries and %s hits gives pi estimate of %s.", tries, hits, pi)
        if output_uri is not None:
            df = spark.createDataFrame([(tries, hits, pi)], ["tries", "hits", "pi"])
            df.write.mode("overwrite").json(output_uri)


if __name__ == "__main__":
    parser = argparse.ArgumentParser()
    parser.add_argument(
        "--partitions",
        default=2,
        type=int,
        help="The number of parallel partitions to use when calculating pi.",
    )
    parser.add_argument(
        "--output_uri", help="The URI where output is saved, typically an S3 bucket."
    )
    args = parser.parse_args()

    calculate_pi(args.partitions, args.output_uri)
```

Amazon S3를 사용하여 Spark 성능 개선

Amazon EMR은 Amazon S3에 저장된 데이터를 쿼리, 읽기 및 쓰기 위해 Spark를 사용할 때 성능을 최적화하는 기능을 제공합니다.

S3 Select는 Amazon S3로 처리를 '푸시다운'하여 일부 애플리케이션의 CSV 및 JSON 파일에 대한 쿼리 성능을 향상시킬 수 있습니다.

EMRFS S3-optimized 커미터는 OutputCommitter 클래스의 대안으로, Spark, DataFrames 및 Datasets를 사용하여 Amazon S3에 Parquet 파일을 쓸 때 EMRFS의 멀티파트 업로드 기능을 사용하여 성능을 개선합니다.

주제

S3 Select와 함께 Spark를 사용하여 쿼리 성능 향상

> **Important**
>
> 신규 고객은 더 이상 Amazon S3 Select를 사용할 수 없습니다. Amazon S3 Select 기존 고객은 평소처럼 이 기능을 계속 사용할 수 있습니다. 자세히 알아보기

Amazon EMR 릴리스 5.17.0 이상에서는 Amazon EMR 기반 Spark에서 S3 Select를 사용할 수 있습니다. S3 Select를 사용하면 애플리케이션이 객체에서 데이터 하위 집합만 검색할 수 있습니다. Amazon EMR의 경우 처리를 위해 대형 데이터 세트를 필터링하는 계산 작업이 클러스터에서 Amazon S3로 '푸시다운'되어 일부 애플리케이션의 성능이 향상되고 Amazon EMR과 Amazon S3 사이에 전송되는 데이터의 양이 줄어듭니다.

S3 Select는 `s3selectCSV` 및 `s3selectJSON` 값을 사용하여 데이터 형식을 지정하는 CSV 및 JSON 파일에서 지원됩니다. 자세한 정보와 지침은 코드에서 S3 Select 지정 섹션을 참조하세요.

S3 Select가 애플리케이션에 적합한가요?

S3 Select 사용 여부에 관계없이 애플리케이션을 벤치마킹하여 애플리케이션에 적합한지 여부를 확인하는 것이 좋습니다.

다음 지침을 사용하여 애플리케이션과 함께 S3 Select를 사용할 수 있는지 확인하십시오.

- 쿼리가 원본 데이터 세트 중 반 이상을 필터링합니다.
- Amazon S3 및 Amazon EMR 클러스터 간 네트워크 연결의 전송 속도와 가용 대역폭이 양호합니다. Amazon S3는 HTTP 응답을 압축하지 않으므로 압축된 입력 파일의 경우 응답 크기가 증가할 수 있습니다.

고려 사항 및 제한 사항

- 고객 제공 암호화 키를 사용하는 Amazon S3 서버 측 암호화(SSE-C) 및 클라이언트 측 암호화는 지원되지 않습니다.
- `AllowQuotedRecordDelimiters` 속성이 지원되지 않습니다. 이 속성이 지정되면 쿼리가 실패합니다.
- UTF-8 형식의 CSV 및 JSON 파일만 지원됩니다. 여러 줄의 CSV는 지원되지 않습니다.
- 비압축 또는 gzip 파일만 지원됩니다.
- `nanValue`, `positiveInf`, `negativeInf` 및 손상된 레코드와 관련된 옵션(예: failfast 및 dropmalformed 모드)은 지원되지 않습니다.
- 10진수 안에 쉼표(,)를 사용하는 것은 지원되지 않습니다. 예를 들어 `10,000`은 지원되지 않고 `10000`은 지원됩니다.
- 마지막 줄의 설명 문자는 지원되지 않습니다.
- 파일 끝의 빈 줄은 처리되지 않습니다.
- 다음과 같은 필터는 Amazon S3로 푸시다운되지 않습니다.
  - `COUNT()` 및 `SUM()` 등의 집계 함수
  - 속성에 대해 `CAST()`를 사용하는 필터링 예를 들어 `CAST(stringColumn as INT) = 1`입니다.
  - 객체이거나 복잡한 속성을 가진 필터 예를 들어 `intArray[1] = 1, objectColumn.objectNumber = 1`입니다.

- 값이 리터럴 값이 아닌 필터 예: `intColumn1 = intColumn2`
- 문서화된 제한 사항으로 S3 Select 지원 데이터 형식만 지원됩니다.

코드에서 S3 Select 지정

다음 예제는 Scala, SQL, R 및 PySpark를 사용하여 CSV에 대해 S3 Select를 지정하는 방법을 보여줍니다. JSON에 대해 S3 Select를 동일한 방식으로 사용할 수 있습니다. 옵션 목록, 기본값 및 제한 사항을 보려면 옵션 단원을 참조하십시오.

PySpark

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional
  .load("s3://path/to/my/datafiles")
```

R

```
read.df("s3://path/to/my/datafiles", "s3selectCSV", schema, header = "true",
 delimiter = "\t")
```

Scala

```
spark
  .read
  .format("s3selectCSV") // "s3selectJson" for Json
  .schema(...) // optional, but recommended
  .options(...) // optional. Examples:
  // .options(Map("quote" -> "\'", "header" -> "true")) or
  // .option("quote", "\'").option("header", "true")
  .load("s3://path/to/my/datafiles")
```

SQL

```
CREATE TEMPORARY VIEW MyView (number INT, name STRING) USING s3selectCSV OPTIONS
 (path "s3://path/to/my/datafiles", header "true", delimiter "\t")
```

옵션

s3selectCSV 및 s3selectJSON을 사용하는 경우 다음 옵션을 사용할 수 있습니다. 지정되지 않은 경우 기본값이 사용됩니다.

S3selectCSV와 옵션

옵션	Default	사용법
compression	"none"	압축 사용 여부를 나타냅니다. "none" 외에도 "gzip"만 지원됩니다.
delimiter	","	필드 구분 기호를 지정합니다.
quote	'\"'	따옴표 문자를 지정합니다. 빈 문자열 지정은 지원되지 않으며 잘못된 XML 오류가 발생합니다.
escape	'\\'	이스케이프 문자를 지정합니다.
header	"false"	"false" 는 헤더가 없음을 나타냅니다. "true"는 헤더가 첫 번째 줄에 있음을 나타냅니다. 첫 번째 줄의 헤더만 지원되며 헤더 앞의 빈 줄은 지원되지 않습니다.
설명	"#"	설명 문자를 지정합니다. 설명 표시기는 비활성화할 수 없습니다. 즉, \u0000의 값은 지원되지 않습니다.
nullValue	""	

selectJSON과 옵션

옵션	Default	사용법
`compression`	`"none"`	압축 사용 여부를 나타냅니다. `"none"` 외에도 `"gzip"`만 지원됩니다.
`multiline`	"false"	`"false"`는 JSON이 S3 Select `LINES` 형식인 것으로 지정합니다. 즉, 입력 데이터의 각 줄에 단일 JSON 객체가 들어 있음을 의미합니다. `"true"`는 JSON이 S3 Select `DOCUMENT` 형식인 것으로 지정합니다. 이는 JSON 객체가 입력 데이터의 여러 줄에 걸쳐 있을 수 있음을 의미합니다.

EMR Spark MagicCommitProtocol

EMR 6.15.0부터는 S3A 파일 시스템을 사용할 때 MagicCommitProtocol이 Spark용 기본 FileCommitProtocol이 됩니다.

MagicCommitProtocol

MagicCommitProtocol은 S3A 파일 시스템을 사용할 때 Amazon S3에 EMR Spark를 사용하여 파일을 쓰는 데 최적화된 FileCommitProtocol의 대체 구현입니다. 이 프로토콜은 작업 및 작업 커밋 단계에서 Amazon S3의 이름 바꾸기 작업을 사용하지 않음으로써 애플리케이션 성능을 개선하는 것을 목표로 합니다.

MagicCommitProtocol은 S3A 파일 시스템을 사용할 때 Amazon Elastic Map Reduce(EMR)에서 실행되는 Spark에서 사용하는 기본 FileCommitProtocol 구현입니다. MagicCommitProtocol은 내부적으로 MagicV2Committer를 사용하여 Amazon S3에 파일 쓰기를 수행합니다.

정적 삽입 작업의 경우 MagicCommitProtocol은 작업 커밋 단계 중에 작업의 출력 위치에 파일을 기록합니다. 반대로 동적 삽입 덮어쓰기 작업의 경우 작업 시도로 작성된 파일은 작업 커밋 시 작업의 출력

위치에만 나타납니다. 이는 작업 커밋 호출 시 커밋 메타데이터를 Spark 드라이버로 다시 내보내면 가능합니다.

MagicCommitProtocol 활성화

MagicCommitProtocol은 S3A 파일 시스템을 사용할 때 Amazon Elastic Map Reduce(EMR)에서 실행되는 Spark에 대해 기본적으로 활성화됩니다.

S3A 파일 시스템을 사용하려면 다음 중 하나를 수행할 수 있습니다.

1. 테이블, 파티션 또는 디렉터리를 정의할 `s3a://` 때 파일 체계를 로 사용합니다.
2. core-site.xml`fs.s3.impl=org.apache.hadoop.fs.s3a.S3AFileSystem`에서 구성을 설정합니다.

MagicCommitProtocol 비활성화

1. 에서 하드 코딩하여 Spark 셸 또는 `spark-submit` `spark-sql` 도구 `SparkConf`또는에서 `--conf` 파라미터로 전달하여 false`spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol.leverageMagic` 설정할 수 있습니다`conf/spark-defaults.conf`. 자세한 내용은 Apache [Spark 설명서의 Spark 구성을](#) 참조하세요.

   다음 예제에서는 `spark-sql` 명령을 실행하는 동안 MagicCommitProtocol을 비활성화하는 방법을 보여줍니다.

   ```
   spark-sql \
     --conf
    spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol.leverageMagicCommitProtocol=fa
    \
   -e "INSERT OVERWRITE TABLE target_table SELECT * FROM source_table;"
   ```

2. `spark-defaults` 구성 분류를 사용하여 `spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol.leverageMagicCom` 속성을 false로 설정합니다. 자세한 내용은 [애플리케이션 구성](#)을 참조하세요.

MagicCommitProtocol 고려 사항

- 정적 파티션 삽입의 경우, Spark 실행기에서 MagicCommitProtocol은 작업이 커밋되거나 중단될 때까지 작업 시도로 작성된 각 파일에 대해 소량의 메모리를 사용합니다. 대부분의 작업에서 사용되는 메모리 양은 무시할 수 있습니다. Spark 드라이버에는 추가 메모리 요구 사항이 없습니다.
- 동적 파티션 삽입의 경우 Spark 드라이버에서 MagicCommitProtocol은 작업이 커밋되거나 중단될 때까지 각 커밋된 파일의 메타데이터 정보를 저장하는 메모리가 필요합니다. 대부분의 작업에서 기본 Spark 드라이버 메모리 설정은 무시해도 됩니다.

  많은 수의 파일을 작성하는 장기 실행 작업이 있는 작업의 경우 커밋 프로토콜이 소비하는 메모리가 눈에 띄고 Spark(특히 Spark 실행기)에 할당된 메모리를 조정해야 할 수도 있습니다. Spark 드라이버의 `spark.driver.memory` 속성과 Spark 실행기의 `spark.executor.memory` 속성을 사용하여 메모리를 튜닝할 수 있습니다. 참고로 100,000개의 파일을 쓰는 단일 작업에는 일반적으로 200MB의 메모리가 추가로 필요합니다. 자세한 내용은 Apache Spark Configuration 설명서에서 Application properties를 참조하세요.

EMRFS S3 최적화 커미터 사용

EMRFS S3 최적화된 커미터는 EMRFS를 사용할 때 Amazon S3에 파일을 쓰는 데 최적화된 대체 OutputCommitter 구현입니다. EMRFS S3 최적화 커미터는 작업 커밋 단계 중 Amazon S3에서 수행된 나열 및 이름 바꾸기 조작을 피함으로써 애플리케이션 성능을 향상시킵니다. 커미터는 Amazon EMR 릴리스 버전 5.19.0 이상에서 사용 가능하며 Amazon EMR 5.20.0 이상에서 기본적으로 활성화됩니다. 커미터는 Spark, DataFrames 또는 데이터 세트를 사용하는 Spark 작업에 사용됩니다. Amazon EMR 6.4.0부터 이 커미터는 Parquet, ORC 및 텍스트 기반 형식(CSV 및 JSON 포함)을 비롯한 모든 일반적인 형식에 사용할 수 있습니다. Amazon EMR 6.4.0 이전 릴리스의 경우 Parquet 형식만 지원됩니다. 커미터가 사용되지 않는 상황이 있습니다. 자세한 내용은 EMRFS S3 최적화 커미터의 요구 사항 단원을 참조하십시오.

주제

EMRFS S3 최적화 커미터의 요구 사항

EMRFS S3 최적화 커미터는 다음과 같은 상황에서 사용됩니다.

- Spark, DataFrames 데이터세트를 사용하여 Amazon S3에 파일을 쓰는 Spark 작업을 실행합니다. Amazon EMR 6.4.0부터 이 커미터는 Parquet, ORC 및 텍스트 기반 형식(CSV 및 JSON 포함)을 비롯한 모든 일반적인 형식에 사용할 수 있습니다. Amazon EMR 6.4.0 이전 릴리스의 경우 Parquet 형식만 지원됩니다.
- Amazon EMR에서 멀티파트 업로드는 활성화됩니다. 이 값이 기본값입니다. 자세한 내용은 [EMRFS S3 최적 커미터 및 멀티파트 업로드](#) 단원을 참조하십시오.
- Spark의 기본 제공 파일 형식 지원이 사용됩니다. 기본 제공 파일 형식 지원은 다음 상황에서 사용됩니다.
  - Hive 메타스토어 테이블의 경우 Parquet 테이블에 대해 `spark.sql.hive.convertMetastoreParquet`을 `true`로 설정하거나, Amazon EMR 6.4.0 이상에서 Orc 테이블에 대해 `spark.sql.hive.convertMetastoreOrc`를 `true`로 설정하는 경우. 기본 설정입니다.
  - 예를 들어, 작업을 파일 형식 데이터 소스 또는 테이블에 쓰는 경우 `USING parquet` 절과 함께 대상 테이블이 생성됩니다.
  - 작업이 파티션 분할되지 않은 Hive 메타스토어 Parquet 테이블에 쓸 경우. Spark의 기본 제공 Parquet 지원은 파티션 분할된 Hive 테이블을 지원하지 않습니다. 이 제한 사항은 알려져 있습니다. 자세한 내용은 Apache Spark, DataFrames 및 데이터 세트 안내서의 [Hive 메타스토어 Parquet 테이블 변환](#)을 참조하세요.
- 기본 파티션 위치(예: `${table_location}/k1=v1/k2=v2/`)에 쓰는 Spark 작업 작업에서는 커미터를 사용합니다. 작업이 사용자 지정 파티션 위치에 쓰는 경우(예: `ALTER TABLE SQL` 명령을 사용하여 사용자 지정 파티션 위치를 설정하는 경우) 커미터는 사용되지 않습니다.
- Spark에는 다음 값을 사용해야 합니다.
  - `spark.sql.parquet.fs.optimized.committer.optimization-enabled` 속성을 `true`로 설정해야 합니다. 이 설정은 Amazon EMR 5.20.0 이상에서 기본 설정입니다. Amazon EMR 5.19.0에서 기본값은 `false`입니다. 이 값의 구성에 대한 자세한 내용은 다음([Amazon EMR 5.19.0에서 EMRFS S3 최적화 커미터 활성화](#))을 참조하십시오.
  - 파티셔닝되지 않은 Hive 메타스토어 테이블에 쓰는 경우 Parquet 및 Orc 파일 형식만 지원됩니다. 파티셔닝되지 않은 Parquet Hive 메타스토어 테이블에 쓸 경우 `spark.sql.hive.convertMetastoreParquet`을 `true`로 설정해야 합니다. 파티셔닝되지 않은 Orc Hive 메타스토어 테이블에 쓸 경우 `spark.sql.hive.convertMetastoreOrc`를 `true`로 설정해야 합니다. 기본 설정입니다.

- spark.sql.parquet.output.committer.class를 com.amazon.emr.committer.EmrOptimizedSparkSqlParquetOutputCommitter로 설정해야 합니다. 이것이 기본 설정입니다.
- spark.sql.sources.commitProtocolClass를 org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol 또는 org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol로 설정해야 합니다. Amazon EMR 5.x 시리즈 버전 5.30.0 이상 및 Amazon EMR 6.x 시리즈 버전 6.2.0 이상에서 org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol이 기본 설정입니다. 이전 Amazon EMR 버전에서는 org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol이 기본 설정입니다.
- Spark 작업이 동적 파티션 열을 포함하는 파티션 분할된 Parquet 데이터 세트를 덮어쓸 경우 partitionOverwriteMode 쓰기 옵션과 spark.sql.sources.partitionOverwriteMode를 static으로 설정해야 합니다. 이것이 기본 설정입니다.

> **Note**
>
> partitionOverwriteMode 쓰기 옵션은 Spark 2.4.0에서 소개되었습니다. Amazon EMR 릴리스 5.19.0에 포함된 Spark 버전 2.3.2의 경우 spark.sql.sources.partitionOverwriteMode 속성을 설정합니다.

EMRFS S3 최적화 커미터를 사용하지 않는 경우

일반적으로 EMRFS S3 최적화 커미터를 다음 상황에서 사용하지 않습니다.

상황	커미터가 사용되지 않는 이유
HDFS에 기록하는 경우	커미터는 EMRFS를 사용하는 Amazon S3에 대한 쓰기만 지원합니다.
S3A 파일 시스템을 사용하는 경우	커미터는 EMRFS만 지원합니다.

상황	커미터가 사용되지 않는 이유
MapReduce 또는 Spark의 RDD API를 사용하는 경우	커미터는 SparkSQL, DataFrame 또는 데이터세트 API 사용만 지원합니다.

다음 Scala 예제에서는 EMRFS S3 최적화 커미터의 사용을 전체적으로(첫 번째 예제) 또는 부분적으로(두 번째 예제) 차단하는 몇 가지 추가 상황을 보여줍니다.

Example - 동적 파티션 덮어쓰기 모드

다음 Scala 예제에서는 Spark에 다른 커밋 알고리즘을 사용하도록 지시합니다. 그러면 EMRFS S3 최적화된 커미터도 사용하지 못합니다. 코드는 데이터를 쓰는 해당 파티션만 덮어쓰도록 `partitionOverwriteMode` 속성을 `dynamic`으로 설정합니다. 그러면 동적 파티션 열이 `partitionBy`에 의해 지정되고 쓰기 모드가 `overwrite`로 설정됩니다.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")
  .option("partitionOverwriteMode", "dynamic")
  .partitionBy("dt")
  .parquet("s3://amzn-s3-demo-bucket1/output")
```

EMRFS S3 최적화 커미터를 사용하지 않도록 세 가지 설정을 모두 구성해야 합니다. 이렇게 하면 Spark는 Spark의 커밋 프로토콜에 지정된 다른 커밋 알고리즘을 실행합니다. 5.30.0 이전의 Amazon EMR 5.x 릴리스와 6.2.0 이전의 Amazon EMR 6.x 릴리스에서 커밋 프로토콜은 Spark의 스테이징 디렉터리를 사용합니다. 이 디렉터리는 `.spark-staging`으로 시작하는 출력 위치 아래에 생성된 임시 디렉터리입니다. 알고리즘은 파티션 디렉터리의 이름을 순차적으로 변경하여 성능을 떨어뜨릴 수 있습니다. Amazon EMR 릴리스 5.30.0 이상과 6.2.0 이상에 대한 자세한 내용은 [EMRFS S3 최적화된 커밋 프로토콜 사용](#) 섹션을 참조하세요.

Spark 2.4.0의 알고리즘은 다음 단계를 수행합니다.

1. 작업 시도에서 출력을 Spark의 스테이징 디렉터리 아래의 파티션 디렉터리(예: `${outputLocation}/spark-staging-${jobID}/k1=v1/k2=v2/`)에 작성합니다.
2. 작성된 각 파티션에 대해 작업 시도에서 상대 파티션 경로(예: `k1=v1/k2=v2`)를 추적합니다.
3. 작업이 성공적으로 완료되면 추적된 모든 상대적 파티션 경로를 드라이버에 제공합니다.

4. 모든 작업이 완료된 후 작업 커밋 단계에서 성공적인 작업 시도가 Spark의 스테이징 디렉터리에 쓴 모든 파티션 디렉터리를 수집합니다. Spark는 디렉터리 트리 이름 변경 작업을 사용하여 각 디렉터리의 이름을 순차적으로 최종 출력 위치로 변경합니다.
5. 스테이징 디렉터리는 작업 커밋 단계가 완료되기 전에 삭제됩니다.

Example - 사용자 지정 파티션 위치

이 예에서 Scala 코드는 두 개의 파티션으로 삽입됩니다. 한 파티션은 사용자 지정 파티션 위치를 갖습니다. 다른 파티션은 기본 파티션 위치를 사용합니다. EMRFS S3 최적화 커미터는 기본 파티션 위치를 사용하는 파티션에 작업 출력을 쓸 때만 사용됩니다.

```
val table = "dataset"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .write.insertInto(table)
```

Scala 코드는 다음 Amazon S3 객체를 생성합니다.

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
```

```
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

사용자 지정 위치의 파티션에 쓸 때 Spark는 앞의 예와 비슷한 커밋 알고리즘을 사용합니다. 이에 대해서는 아래에서 간단히 설명합니다. 앞의 예제와 같이 이 알고리즘은 이름을 순차적으로 변경하여 성능에 부정적인 영향을 줄 수 있습니다.

1. 사용자 지정 위치의 파티션에 출력을 쓸 경우 작업은 최종 출력 위치에 생성된 Spark의 스테이징 디렉터리에 있는 파일에 씁니다. 파일 이름에는 파일 충돌을 방지하기 위해 무작위 UUID가 포함됩니다. 작업 시도는 각 파일의 추적과, 필요한 최종 출력 경로를 보존합니다.
2. 작업이 성공적으로 완료되면 드라이버에 파일과 해당 출력 경로를 제공합니다.
3. 모든 작업이 완료된 후 작업 커밋 단계에서는 사용자 지정 위치에 파티션에 대해 써진 모든 파일의 이름을 최종 출력 경로로 순차적으로 변경합니다.
4. 스테이징 디렉터리는 작업 커밋 단계가 완료되기 전에 삭제됩니다.

EMRFS S3 최적 커미터 및 멀티파트 업로드

EMRFS S3 최적화 커미터를 사용하려면 Amazon EMR에 대한 멀티파트 업로드를 활성화해야 합니다. 멀티파트 업로드는 기본적으로 사용하도록 설정되어 있습니다. 필요한 경우 다시 활성화할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서의 Amazon S3에 대한 멀티파트 업로드 구성을 참조하세요.

EMRFS S3 최적화 커미터는 멀티파트 업로드의 트랜잭션과 유사한 특성을 사용하여 작업을 시도할 때 작성된 파일이 작업 커밋 시 작업 출력 위치에만 나타나는지 확인합니다. 이 방법으로 멀티파트 업로드를 사용하면 커미터는 기본 FileOutputCommitter 알고리즘 버전 2보다 작업 커밋 성능을 향상시킵니다. EMRFS S3 최적화 커미터를 사용할 경우, 일반적인 멀티파트 업로드 동작과의 몇 가지 주요한 차이점을 고려해야 합니다.

- 멀티파트 업로드는 파일 크기에 관계없이 항상 수행됩니다. 이는 `fs.s3n.multipart.uploads.split.size` 속성이 멀티파트 업로드가 트리거되는 파일 크기를 제어하는 EMRFS의 기본 동작과 다릅니다.
- 멀티파트 업로드는 작업이 커밋되거나 중단될 때까지 더 오랜 기간 동안 불완전한 상태로 유지됩니다. 이는 지정된 파일 작성을 마칠 때 멀티파트 업로드가 완료되는 EMRFS의 기본 동작과 다릅니다.

이러한 차이로 인해 Spark Executor JVM이 충돌하거나 작업이 실행 중이고 Amazon S3에 데이터를 쓰는 경우 불완전한 멀티파트 업로드가 남아 있게 됩니다. 이러한 이유로 EMRFS S3 최적화 커미터를 사용할 때는 실패한 멀티파트 업로드를 관리하는 모범 사례를 따르십시오. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon S3 버킷 사용에 대한 모범 사례를 참조하세요.

작업 튜닝 고려 사항

EMRFS S3 최적화 커미터는 작업이 커밋되거나 중단될 때까지 작업 시도 시 작성된 각 파일에 대해 소량의 메모리를 사용합니다. 대부분의 작업에서 사용되는 메모리 양은 무시할 수 있습니다. 많은 수의 파일을 작성하는 장기 실행 작업이 있는 작업의 경우 커미터가 소비하는 메모리가 눈에 띄고 Spark 실행기에 할당된 메모리를 조정해야 할 수도 있습니다. `spark.executor.memory` 속성을 사용하여 실행기 메모리를 튜닝할 수 있습니다. 참고로, 100,000개의 파일을 작성하는 단일 작업에는 일반적으로 100MB의 메모리가 추가로 필요합니다. 자세한 내용은 Apache Spark Configuration 설명서에서 Application properties를 참조하세요.

Amazon EMR 5.19.0에서 EMRFS S3 최적화 커미터 활성화

Amazon EMR 5.19.0을 사용할 경우 클러스터를 만들 때 `spark.sql.parquet.fs.optimized.committer.optimization-enabled` 속성을 `true`로 직접 설정할 수 있습니다. Amazon EMR을 사용할 때는 Spark에서 설정할 수 있습니다.

클러스터 생성 시 EMRFS S3 최적화 커미터 활성화

`spark-defaults` 구성 분류를 사용하여 `spark.sql.parquet.fs.optimized.committer.optimization-enabled` 속성을 `true`로 설정합니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

Spark에서 EMRFS S3 최적화 커미터 활성화

`SparkConf`로 하드 코딩하여 `spark.sql.parquet.fs.optimized.committer.optimization-enabled`를 `true`로 설정하고 Spark 셸의 `--conf` 파라미터 또는 `spark-submit` 및 `spark-sql` 도구 또는 `conf/spark-defaults.conf`로 전달할 수 있습니다. 자세한 내용은 Apache Spark 설명서에서 Spark configuration을 참조하세요.

다음 예제에서는 spark-sql 명령을 실행하는 동안 커미터를 활성화하는 방법을 보여줍니다.

```
spark-sql \
  --conf spark.sql.parquet.fs.optimized.committer.optimization-enabled=true \
  -e "INSERT OVERWRITE TABLE target_table SELECT * FROM source_table;"
```

EMRFS S3 최적화된 커밋 프로토콜 사용

EMRFS S3 최적화된 커밋 프로토콜은 EMRFS를 사용할 때 Spark 동적 파티션 덮어쓰기로 Amazon S3에 파일을 쓰는 데 최적화된 대체 FileCommitProtocol 구현입니다. 이 프로토콜은 Spark 동적 파티션 덮어쓰기 작업 커밋 단계 중 Amazon S3에서 이름 바꾸기 작업을 피함으로써 애플리케이션 성능을 개선합니다.

이름 바꾸기 작업을 피함으로써 EMRFS S3 최적화 커미터 사용 성능도 향상됩니다. 하지만 커밋 프로토콜의 개선 사항이 동적 파티션 덮어쓰기 사례만 대상으로 하는 반면 이 기능은 동적 파티션 덮어쓰기 사례에 유효하지 않습니다.

커밋 프로토콜은 Amazon EMR 릴리스 버전 5.30.0 이상 및 6.2.0 이상에서 사용 가능하며 기본적으로 사용 가능합니다. Amazon EMR은 릴리스 5.31.0부터 병렬 처리 개선 기능을 추가했습니다. 프로토콜은 Spark, DataFrames 데이터 세트를 사용하는 Spark 작업에 사용됩니다. 커밋 프로토콜이 사용되지 않는 상황이 있습니다. 자세한 내용은 EMRFS S3 최적화 커밋 프로토콜의 요구 사항 단원을 참조하십시오.

주제

EMRFS S3 최적화 커밋 프로토콜의 요구 사항

EMRFS S3 최적화 커밋 프로토콜은 다음 조건을 충족하는 경우에 사용됩니다.

- Spark, DataFrames 데이터 세트를 사용하여 분할된 테이블을 덮어쓰는 Spark 작업을 실행합니다.
- 파티션 덮어쓰기 모드가 `dynamic`인 Spark 작업을 실행합니다.
- Amazon EMR에서 멀티파트 업로드는 활성화됩니다. 이 값이 기본값입니다. 자세한 내용은 EMRFS S3 최적화 커밋 프로토콜 및 멀티파트 업로드 단원을 참조하십시오.
- EMRFS의 파일 시스템 캐시가 활성화되었습니다. 이 값이 기본값입니다. `fs.s3.impl.disable.cache` 설정이 `false`로 설정되어 있는지 확인합니다.
- Spark의 기본 제공 데이터 소스 지원이 사용됩니다. 기본 제공 데이터 소스 지원은 다음 상황에서 사용됩니다.
  - 작업에서 기본 제공 데이터 소스 또는 테이블에 쓰는 경우.

  - 작업에서 Hive 메타스토어 Parquet 테이블에 쓰는 경우. `spark.sql.hive.convertInsertingPartitionedTable` 및 `spark.sql.hive.convertMetastoreParquet` 모두 true로 설정된 경우에 수행됩니다. 기본 설정입니다.
  - 작업에서 Hive 메타스토어 ORC 테이블에 쓰는 경우. `spark.sql.hive.convertInsertingPartitionedTable` 및 `spark.sql.hive.convertMetastoreOrc` 모두 `true`로 설정된 경우에 수행됩니다. 기본 설정입니다.
- 기본 파티션 위치에 쓰는 Spark 작업(예: `${table_location}/k1=v1/k2=v2/`)에서는 커밋 프로토콜을 사용합니다. 작업이 사용자 지정 파티션 위치에 쓰는 경우(예: `ALTER TABLE SQL` 명령을 사용하여 사용자 지정 파티션 위치를 설정하는 경우) 프로토콜은 사용되지 않습니다.
- Spark에는 다음 값을 사용해야 합니다.
  - `spark.sql.sources.commitProtocolClass`를 `org.apache.spark.sql.execution.datasources.SQLEmrOptimizedCommitProtocol`로 설정해야 합니다. 이는 Amazon EMR 릴리스 5.30.0 이상 및 6.2.0 이상 릴리스에서 기본 설정입니다.
  - `partitionOverwriteMode` 쓰기 옵션 또는 `spark.sql.sources.partitionOverwriteMode`를 `dynamic`으로 설정해야 합니다. 기본 설정은 `static`입니다.

    > **Note**
    >
    > `partitionOverwriteMode` 쓰기 옵션은 Spark 2.4.0에서 소개되었습니다. Amazon EMR 릴리스 5.19.0에 포함된 Spark 버전 2.3.2의 경우 `spark.sql.sources.partitionOverwriteMode` 속성을 설정합니다.

  - Spark 작업이 Hive 메타스토어 Parquet 테이블을 덮어쓰는 경우 `spark.sql.hive.convertMetastoreParquet`, `spark.sql.hive.convertInsertingPartitionedTable` 및 `spark.sql.hive.convertMetastore.partitionOverwriteMode`를 `true`로 설정해야 합니다. 기본 설정이 있습니다.
  - Spark 작업이 Hive 메타스토어 ORC 테이블을 덮어쓰는 경우 `spark.sql.hive.convertMetastoreOrc`, `spark.sql.hive.convertInsertingPartitionedTable` 및

spark.sql.hive.convertMetastore.partitionOverwriteMode를 true로 설정해야 합니다. 기본 설정이 있습니다.

Example - 동적 파티션 덮어쓰기 모드

이 Scala 예제에서는 최적화가 트리거됩니다. 먼저 partitionOverwriteMode 속성을 dynamic으로 설정합니다. 이렇게 하면 데이터를 쓰는 대상 파티션만 덮어씁니다. 그런 다음, partitionBy를 사용하여 동적 파티션 열을 지정하고 쓰기 모드를 overwrite로 설정합니다.

```
val dataset = spark.range(0, 10)
  .withColumn("dt", expr("date_sub(current_date(), id)"))

dataset.write.mode("overwrite")                 // "overwrite" instead of "insert"
  .option("partitionOverwriteMode", "dynamic")  // "dynamic" instead of "static"
  .partitionBy("dt")                            // partitioned data instead of
 unpartitioned data
  .parquet("s3://amzn-s3-demo-bucket1/output")    // "s3://" to use Amazon EMR file
 system, instead of "s3a://" or "hdfs://"
```

EMRFS S3 최적화 커밋 프로토콜을 사용하지 않는 경우

일반적으로 EMRFS S3-optimized 커밋 프로토콜은 오픈 소스 기본 Spark 커밋 프로토콜인와 동일하게 작동합니다org.apache.spark.sql.execution.datasources.SQLHadoopMapReduceCommitProtocol. 다음과 같은 상황에서는 최적화가 수행되지 않습니다.

상황	커밋 프로토콜이 사용되지 않는 이유
HDFS에 기록하는 경우	커밋 프로토콜은 EMRFS를 사용하는 Amazon S3에 대한 쓰기만 지원합니다.
S3A 파일 시스템을 사용하는 경우	커밋 프로토콜은 EMRFS만 지원합니다.
MapReduce 또는 Spark의 RDD API를 사용하는 경우	커밋 프로토콜은 SparkSQL, DataFrame 또는 데이터 세트 API 사용만 지원합니다.
동적 파티션 덮어쓰기가 트리거되지 않는 경우	커밋 프로토콜은 동적 파티션 덮어쓰기 사례만 최적화합니다. 그 밖의 사례는 EMRFS S3 최적화 커미터 사용 섹션을 참조하세요.

다음 Scala 예제에서는 EMRFS S3 최적화 커밋 프로토콜이 `SQLHadoopMapReduceCommitProtocol`로 위임하는 몇 가지 추가 상황을 보여줍니다.

Example - 사용자 지정 파티션 위치가 있는 동적 파티션 덮어쓰기 모드

이 예제에서 Scala 프로그램은 동적 파티션 덮어쓰기 모드에서 두 개의 파티션을 덮어씁니다. 한 파티션은 사용자 지정 파티션 위치를 갖습니다. 다른 파티션은 기본 파티션 위치를 사용합니다. EMRFS S3 최적화 커미터는 기본 파티션 위치를 사용하는 파티션만 개선합니다.

```
val table = "dataset"
val inputView = "tempView"
val location = "s3://bucket/table"

spark.sql(s"""
  CREATE TABLE $table (id bigint, dt date)
  USING PARQUET PARTITIONED BY (dt)
  LOCATION '$location'
""")

// Add a partition using a custom location
val customPartitionLocation = "s3://bucket/custom"
spark.sql(s"""
  ALTER TABLE $table ADD PARTITION (dt='2019-01-28')
  LOCATION '$customPartitionLocation'
""")

// Add another partition using default location
spark.sql(s"ALTER TABLE $table ADD PARTITION (dt='2019-01-29')")

def asDate(text: String) = lit(text).cast("date")

spark.range(0, 10)
  .withColumn("dt",
    when($"id" > 4, asDate("2019-01-28")).otherwise(asDate("2019-01-29")))
  .createTempView(inputView)

// Set partition overwrite mode to 'dynamic'
spark.sql(s"SET spark.sql.sources.partitionOverwriteMode=dynamic")

spark.sql(s"INSERT OVERWRITE TABLE $table SELECT * FROM $inputView")
```

Scala 코드는 다음 Amazon S3 객체를 생성합니다.

```
custom/part-00001-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
custom_$folder$
table/_SUCCESS
table/dt=2019-01-29/part-00000-035a2a9c-4a09-4917-8819-e77134342402.c000.snappy.parquet
table/dt=2019-01-29_$folder$
table_$folder$
```

> **Note**
>
> 이전 Spark 버전에서 사용자 지정 파티션 위치에 데이터를 쓰면 데이터가 손실될 수 있습니다. 이 예제에서는 `dt='2019-01-28'` 파티션이 손실됩니다. 자세한 내용은 SPARK-35106을 참조하세요. 이 문제는 Amazon EMR 릴리스 5.33.0 이상(6.0.x 및 6.1.x 제외)에서 수정되었습니다 .

사용자 지정 위치의 파티션에 쓸 때 Spark는 앞의 예와 비슷한 커밋 알고리즘을 사용합니다. 이에 대해서는 아래에서 간단히 설명합니다. 앞의 예제와 같이 이 알고리즘은 이름을 순차적으로 변경하여 성능에 부정적인 영향을 줄 수 있습니다.

Spark 2.4.0의 알고리즘은 다음 단계를 수행합니다.

1. 사용자 지정 위치의 파티션에 출력을 쓸 경우 작업은 최종 출력 위치에 생성된 Spark의 스테이징 디렉터리에 있는 파일에 씁니다. 파일 이름에는 파일 충돌을 방지하기 위해 무작위 UUID가 포함됩니다. 작업 시도는 각 파일의 추적과, 필요한 최종 출력 경로를 보존합니다.
2. 작업이 성공적으로 완료되면 드라이버에 파일과 해당 출력 경로를 제공합니다.
3. 모든 작업이 완료된 후 작업 커밋 단계에서는 사용자 지정 위치에 파티션에 대해 써진 모든 파일의 이름을 최종 출력 경로로 순차적으로 변경합니다.
4. 스테이징 디렉터리는 작업 커밋 단계가 완료되기 전에 삭제됩니다.

EMRFS S3 최적화 커밋 프로토콜 및 멀티파트 업로드

EMRFS S3 최적화 커밋 프로토콜에서 동적 파티션 덮어쓰기 최적화를 사용하려면 Amazon EMR에서 멀티파트 업로드를 활성화해야 합니다. 멀티파트 업로드는 기본적으로 사용하도록 설정되어 있습니다. 필요한 경우 다시 활성화할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서의 Amazon S3에 대한 멀티파트 업로드 구성을 참조하세요.

동적 파티션 덮어쓰기 중에 EMRFS S3 최적화 커밋 프로토콜은 멀티파트 업로드의 트랜잭션과 유사한 특성을 사용하여 작업을 시도할 때 작성된 파일이 작업 커밋 시 작업 출력 위치에

만 나타나도록 합니다. 이러한 방식으로 멀티파트 업로드를 사용하면 커밋 프로토콜은 기본 `SQLHadoopMapReduceCommitProtocol`보다 작업 커밋 성능을 더 개선합니다. EMRFS S3 최적화 커밋 프로토콜을 사용할 경우, 일반적인 멀티파트 업로드 동작과의 몇 가지 주요한 차이점을 고려해야 합니다.

- 멀티파트 업로드는 파일 크기에 관계없이 항상 수행됩니다. 이는 `fs.s3n.multipart.uploads.split.size` 속성이 멀티파트 업로드가 트리거되는 파일 크기를 제어하는 EMRFS의 기본 동작과 다릅니다.
- 멀티파트 업로드는 작업이 커밋되거나 중단될 때까지 더 오랜 기간 동안 불완전한 상태로 유지됩니다. 이는 지정된 파일 작성을 마칠 때 멀티파트 업로드가 완료되는 EMRFS의 기본 동작과 다릅니다.

이러한 차이로 인해 작업이 실행 중이고 Amazon S3에 데이터를 쓰는 동안 Spark 실행기 JVM이 충돌하거나 강제 종료되거나 작업 실행 중에 Spark 드라이버 JVM이 충돌하거나 강제 종료되는 경우 불완전한 멀티파트 업로드가 남을 가능성이 큽니다. 이러한 이유로 EMRFS S3 최적화 커밋 프로토콜을 사용할 때는 실패한 멀티파트 업로드를 관리하는 모범 사례를 따릅니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon S3 버킷 사용에 대한 모범 사례를 참조하세요.

작업 튜닝 고려 사항

Spark 실행기에서 EMRFS S3 최적화 커밋 프로토콜은 작업이 커밋되거나 중단될 때까지 작업 시도 시 작성된 각 파일에 대해 소량의 메모리를 사용합니다. 대부분의 작업에서 사용되는 메모리 양은 무시할 수 있습니다.

Spark 드라이버에서 EMRFS S3 최적화 커밋 프로토콜에는 작업이 커밋되거나 중단될 때까지 각 커밋된 파일의 메타데이터 정보를 저장하기 위한 메모리가 필요합니다. 대부분의 작업에서 기본 Spark 드라이버 메모리 설정은 무시해도 됩니다.

많은 수의 파일을 작성하는 장기 실행 작업이 있는 작업의 경우 커밋 프로토콜이 소비하는 메모리가 눈에 띄고 Spark(특히 Spark 실행기)에 할당된 메모리를 조정해야 할 수도 있습니다. Spark 드라이버의 `spark.driver.memory` 속성과 Spark 실행기의 `spark.executor.memory` 속성을 사용하여 메모리를 튜닝할 수 있습니다. 참고로, 100,000개의 파일을 작성하는 단일 작업에는 일반적으로 100MB의 메모리가 추가로 필요합니다. 자세한 내용은 Apache Spark Configuration 설명서에서 Application properties를 참조하세요.

EMRFS로 Amazon S3 요청 재시도

이 주제에서는 EMRFS를 사용하여 Amazon S3에 요청할 때 사용할 수 있는 재시도 전략에 대한 정보를 제공합니다. 요청 비율이 증가하면 S3는 새 속도를 지원하도록 조정을 시도합니다. 이 과정에서 S3

는 요청을 제한하고 `503 Slow Down` 오류를 반환할 수 있습니다. S3 요청의 성공률을 높이기 위해 `emrfs-site` 구성에서 속성을 구성하여 재시도 전략을 조정할 수 있습니다.

다음과 같은 방법으로 재시도 전략을 조정할 수 있습니다.

- 기본 지수 백오프 재시도 전략의 최대 재시도 한도를 늘립니다.
- 가산 증가 및 지수 감소(AIMD) 재시도 전략을 활성화하고 구성합니다. AIMD는 Amazon EMR 릴리스 6.4.0 이상에서 지원됩니다.

기본 지수 백오프 전략 사용

기본적으로 EMRFS는 지수 백오프 전략을 사용하여 Amazon S3 요청을 재시도합니다. 기본 EMRFS 재시도 제한은 15회입니다. S3 `503 Slow Down` 오류를 방지하기 위해 새 클러스터를 생성할 때, 실행 중인 클러스터에서 또는 애플리케이션 런타임에 재시도 한도를 늘릴 수 있습니다.

재시도 한도를 늘리려면 `fs.s3.maxRetries` 구성에서 `emrfs-site`의 값을 변경해야 합니다. 다음 예제 구성은 `fs.s3.maxRetries`를 사용자 지정 값(30)으로 설정합니다.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.maxRetries": "30"
      }
    }
]
```

구성 객체 작업에 대한 자세한 내용은 애플리케이션 구성 섹션을 참조하세요.

AIMD 재시도 전략 사용

Amazon EMR 릴리스 6.4.0 이상에서 EMRFS는 가산 증가 및 지수 감소(AIMD) 모델을 기반으로 하는 대체 재시도 전략을 지원합니다. AIMD 재시도 전략은 대형 Amazon EMR 클러스터를 사용할 때 특히 유용합니다.

AIMD는 최근 성공한 요청에 대한 데이터를 사용하여 사용자 지정 요청 비율을 계산합니다. 이 전략을 사용하면 제한된 요청 수와 요청당 필요한 총 시도 횟수를 줄일 수 있습니다.

AIMD 재시도 전략을 활성화하려면 다음 예제와 같이 `emrfs-site` 구성에서 `fs.s3.aimd.enabled` 속성을 `true`로 설정해야 합니다.

```
[
    {
      "Classification": "emrfs-site",
      "Properties": {
        "fs.s3.aimd.enabled": "true"
      }
    }
]
```

구성 객체 작업에 대한 자세한 내용은 애플리케이션 구성 섹션을 참조하세요.

고급 AIMD 재시도 설정

다음 테이블에 나열된 속성을 구성하여 AIMD 재시도 전략을 사용할 때 재시도 동작을 세분화할 수 있습니다. 대부분의 사용 사례에서는 기본값을 사용하는 것이 좋습니다.

고급 AIMD 재시도 전략 속성

속성	기본값	설명
`fs.s3.aimd.increaseIncrement`	0.1	연속 요청 성공 시 요청 비율이 얼마나 빨리 증가하는지를 제어합니다.
`fs.s3.aimd.reductionFactor`	2	Amazon S3에서 503 응답을 반환할 때 요청 비율이 얼마나 빨리 감소하는지를 제어합니다. 기본 계수(2)를 사용하면 요청 비율이 절반으로 줄어듭니다.
`fs.s3.aimd.minRate`	0.1	요청이 S3에 의해 지속적으로 제한되는 경우 요청 비율의 하한을 설정합니다.
`fs.s3.aimd.initialRate`	5500	초기 요청 비율을 설정합니다. 이 속도는 `fs.s3.aimd.increaseIncrement` 및 `fs.s3.aimd.reductionFactor` 에 지정한 값에 따라 달라집니다.

속성	기본값	설명
		초기 속도는 GET 요청에도 사용되며 PUT 요청의 경우 비례적으로 조정됩니다(3,500/5,500).
`fs.s3.aimd.adjustWindow`	2	요청 비율을 조정하는 빈도를 제어합니다(단위: 응답 수).
`fs.s3.aimd.maxAttempts`	100	요청을 시도할 수 있는 최대 시도 횟수를 설정합니다.

Spark 단계 추가

Amazon EMR 단계를 사용하여 EMR 클러스터에 설치된 Spark 프레임워크로 작업을 제출할 수 있습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 단계를 참조하세요. 콘솔 및 CLI에서 Spark 애플리케이션 단계를 사용하여 이 작업을 수행할 수 있습니다. 그러면 사용자를 대신하여 `spark-submit` 스크립트가 실행됩니다. API에서는 단계를 사용하여 `spark-submit` 파일을 통해 `command-runner.jar`를 호출할 수 있습니다.

Spark로 애플리케이션 제출에 대한 자세한 내용은 Apache Spark 설명서에서 Submitting applications 주제를 참조하세요.

콘솔을 사용하여 Spark 단계를 제출하려면

1. https://console.aws.amazon.com/emr Amazon EMR 콘솔을 엽니다.
2. 클러스터 목록에서 클러스터의 이름을 선택합니다.
3. 단계 섹션으로 스크롤하여 섹션을 확장한 후 단계 추가를 선택합니다.
4. 단계 추가 대화 상자에서 다음을 수행합니다.

   - 단계 유형에서 Spark 애플리케이션을 선택합니다.
   - 이름에서 기본 이름(Spark 애플리케이션)을 그대로 사용하거나 새 이름을 입력합니다.
   - Deploy mode(배포 모드)에서 클라이언트 또는 클러스터 모드를 선택합니다. 클라이언트 모드는 클러스터의 프라이머리 인스턴스에서 드라이버 프로그램을 시작하고, 클러스터 모드는 클러스터에서 드라이버 프로그램을 시작합니다. 클라이언트 모드의 경우 드라이버의 로그 출력이 단계 로그에 나타나고, 클러스터 모드의 경우 드라이버의 로그 출력이 첫 번째 YARN 컨테이너

의 로그에 나타납니다. 자세한 내용은 Apache Spark 설명서의 Cluster mode overview를 참조하세요.

- 원하는 Spark-submit 옵션을 지정합니다. `spark-submit` 옵션에 대한 자세한 내용은 Launching applications with spark-submit을 참조하세요.
- 애플리케이션 위치에는 해당 애플리케이션의 로컬 또는 S3 URI 경로를 지정합니다.
- 인수에서 필드를 비워 둡니다.
- 실패 시 작업에서 기본 옵션(계속)을 그대로 사용합니다.

5. 추가를 선택합니다. 단계가 콘솔에 [Pending] 상태로 나타납니다.
6. 단계의 상태는 단계가 실행됨에 따라 대기 중에서 실행 중을 거쳐 완료됨으로 바뀝니다. 상태를 업데이트하려면 작업 열 위의 새로 고침 아이콘을 선택합니다.
7. 단계 결과는 로깅을 구성한 경우 Amazon EMR 콘솔 클러스터 세부 정보 페이지에서 로그 파일 아래 단계 옆에 있습니다. 또는 클러스터를 시작할 때 구성한 로그 버킷에서도 단계 정보를 찾아볼 수 있습니다.

를 사용하여 Spark에 작업을 제출하려면 AWS CLI

클러스터 생성 시 단계를 제출하거나 기존 클러스터에서 `aws emr add-steps` 하위 명령을 사용합니다.

1. 다음 예제에 표시된 대로 `create-cluster`를 사용합니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 --
applications Name=Spark \
--ec2-attributes KeyName=myKey --instance-type m5.xlarge --instance-count 3 \
--steps Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10] --use-default-roles
```

또는 다음 예제에 표시된 대로 `command-runner.jar`을 사용할 수 있습니다.

```
aws emr create-cluster --name "Add Spark Step Cluster" --release-label emr-7.9.0 \
--applications Name=Spark --ec2-attributes KeyName=myKey --instance-type m5.xlarge
 --instance-count 3 \
--steps Type=CUSTOM_JAR,Name="Spark Program",Jar="command-
runner.jar",ActionOnFailure=CONTINUE,Args=[spark-example,SparkPi,10] --use-default-
roles
```

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

2. 또는 이미 실행 중인 클러스터에 단계를 추가합니다. `add-steps`를 사용합니다.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps
 Type=Spark,Name="Spark Program",ActionOnFailure=CONTINUE,Args=[--
class,org.apache.spark.examples.SparkPi,/usr/lib/spark/examples/jars/spark-
examples.jar,10]
```

   또는 다음 예제에 표시된 대로 `command-runner.jar`을 사용할 수 있습니다.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF --steps Type=CUSTOM_JAR,Name="Spark
 Program",Jar="command-runner.jar",ActionOnFailure=CONTINUE,Args=[spark-
example,SparkPi,10]
```

SDK for Java를 사용하여 작업을 Spark로 제출하는 방법

1. 다음은 Java를 사용하여 Spark가 설치된 클러스터에 단계를 추가하는 방법을 보여주는 예제입니다.

```
AWSCredentials credentials = new BasicAWSCredentials(accessKey, secretKey);
AmazonElasticMapReduce emr = new AmazonElasticMapReduceClient(credentials);

StepFactory stepFactory = new StepFactory();
AmazonElasticMapReduceClient emr = new AmazonElasticMapReduceClient(credentials);
AddJobFlowStepsRequest req = new AddJobFlowStepsRequest();
req.withJobFlowId("j-1K48XXXXXXHCB");
```

```
List<StepConfig> stepConfigs = new ArrayList<StepConfig>();

HadoopJarStepConfig sparkStepConf = new HadoopJarStepConfig()
   .withJar("command-runner.jar")
   .withArgs("spark-submit","--executor-memory","1g","--
class","org.apache.spark.examples.SparkPi","/usr/lib/spark/examples/jars/spark-
examples.jar","10");

StepConfig sparkStep = new StepConfig()
   .withName("Spark Step")
   .withActionOnFailure("CONTINUE")
   .withHadoopJarStep(sparkStepConf);

stepConfigs.add(sparkStep);
req.withSteps(stepConfigs);
AddJobFlowStepsResult result = emr.addJobFlowSteps(req);
```

2. 단계에 대한 로그를 검토하여 단계의 결과를 살펴봅니다. 단계를 선택하고 단계를 선택한 다음 로그 파일에서 `stdout` 또는 중 하나를 선택하여 로깅을 활성화한 AWS Management Console 경우에서이 작업을 수행할 수 있습니다`stderr`. 사용할 수 있는 로그를 보려면 로그 보기를 선택합니다.

Spark 기본 구성 설정 재정의

애플리케이션마다 Spark 기본 구성 값을 재정의해야 할 때가 있습니다. 이 작업은 단계를 사용하여 애플리케이션을 제출할 때 기본적으로 옵션을 `spark-submit`에 전달함으로써 수행할 수 있습니다. 예를 들면 `spark.executor.memory`를 변경하여 실행자 프로세스에 할당된 메모리를 변경해야 할 때가 있습니다. 다음과 같은 인수를 사용하여 `--executor-memory`를 지정합니다.

```
spark-submit --executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/
spark/examples/jars/spark-examples.jar 10
```

마찬가지로, `--executor-cores` 및 `--driver-memory`를 조정할 수 있습니다. 단계에서 다음 인수를 단계에 제공합니다.

```
--executor-memory 1g --class org.apache.spark.examples.SparkPi /usr/lib/spark/examples/
jars/spark-examples.jar 10
```

또한 `--conf` 옵션을 사용하여 기본 제공 스위치가 없는 설정을 조정할 수도 있습니다. 조정 가능한 다양한 설정에 대한 자세한 내용은 Apache Spark 설명서의 Dynamically Loading Spark Properties 주제를 참조하세요.

Spark 애플리케이션 기록 보기

콘솔의 클러스터 세부 정보 페이지에서 애플리케이션 사용자 인터페이스 탭을 사용하여 Spark, YARN 애플리케이션 및 Tez UI 세부 정보를 볼 수 있습니다. Amazon EMR 애플리케이션 사용자 인터페이스(UI)를 사용하면 활성 작업 및 작업 내역을 쉽게 분석하고 관련 문제를 해결할 수 있습니다.

자세한 내용은 Amazon EMR 관리 안내서에서 애플리케이션 기록 보기를 참조하세요.

Spark 웹 UI에 액세스

Amazon EMR 관리 안내서에서 클러스터에 연결 섹션에 나온 SSH 터널을 생성하거나 프록시를 생성하는 절차를 수행한 다음, 클러스터의 YARN ResourceManager로 이동하여 Spark 웹 UI를 볼 수 있습니다. 애플리케이션의 Tracking UI(추적 UI)에서 링크를 선택합니다. 애플리케이션이 실행 중인 경우 ApplicationMaster가 표시됩니다. ApplicationMaster는 드라이버가 있는 경우 언제든지 포트 20888의 애플리케이션 마스터 웹 UI로 안내합니다. YARN 클라이언트 모드로 실행하는 경우 드라이버는 클러스터의 프라이머리 노드에 있을 수 있습니다. 애플리케이션을 YARN 클러스터 모드에서 실행 중인 경우 드라이버는 클러스터에서 해당 애플리케이션에 대한 ApplicationMaster에 있습니다. 애플리케이션이 완료된 경우에는 기록이 표시되며, EMR 클러스터 프라이머리 노드의 포트 번호 18080에 있는 Spark HistoryServer UI로 안내합니다. 이 설명은 이미 완료된 애플리케이션에 적용됩니다. 또한 http://*master-public-dns-name*:18080/에서 Spark HistoryServer UI로 직접 이동할 수도 있습니다.

Amazon EMR 릴리스 5.25.0 이상에서는 SSH 연결을 통해 웹 프록시를 설정하지 않고도 콘솔에서 Spark 기록 서버 UI에 액세스할 수 있습니다. 자세한 내용은 영구 애플리케이션 사용자 인터페이스 보기를 참조하세요.

Spark 구조화된 스트리밍 Amazon Kinesis Data Streams 커넥터 사용

Amazon EMR 릴리스 7.1.0 이상에는 릴리스 이미지에 Spark 구조화된 스트리밍 Amazon Kinesis Data Streams 커넥터가 포함되어 있습니다. 이 커넥터를 사용하면 Amazon EMR에서 Spark를 사용하여 Amazon Kinesis Data Streams에 저장된 데이터를 처리할 수 있습니다. 커넥터는 `GetRecords`(공유 처리량) 및 `SubscribeToShard`(향상된 팬아웃)의 소비자 유형을 모두 지원합니다. 이 통합은

spark-sql-kinesis-connector에 기반합니다. 커넥터 사용을 시작하는 방법에 대한 자세한 내용은 README를 참조하세요.

다음 예제에서는 커넥터를 사용하여 Amazon EMR에서 Spark 애플리케이션을 시작하는 방법을 보여 줍니다.

```
spark-submit my_kinesis_streaming_script.py
```

Amazon EMR에서 Apache Spark용 Amazon Redshift 통합 사용

Amazon EMR 릴리스 6.4.0 이상에서 모든 릴리스 이미지에 Apache Spark와 Amazon Redshift 간 커넥터가 포함됩니다. 이 커넥터를 사용하면 Amazon EMR에서 Spark를 사용하여 Amazon Redshift에 저장된 데이터를 처리할 수 있습니다. Amazon EMR 릴리스 6.4.0~6.8.0의 경우 통합은 spark-redshift 오픈 소스 커넥터를 기반으로 합니다. Amazon EMR 릴리스 6.9.0 이상의 경우 Apache Spark용 Amazon Redshift 통합이 커뮤니티 버전에서 네이티브 통합으로 마이그레이션되었습니다.

주제

Apache Spark용 Amazon Redshift 통합을 사용하여 Spark 애플리케이션 시작

Amazon EMR 릴리스 6.4~6.9의 경우 `--jars` 또는 `--packages` 옵션을 사용하여 다음 JAR 파일 중 사용하려는 파일을 지정해야 합니다. `--jars` 옵션은 로컬, HDFS 또는 HTTP 및 HTTPS를 사용하여 저장되는 종속 항목을 지정합니다. `--jars` 옵션에서 지원하는 다른 파일 위치를 보려면 Spark 설명서에서 Advanced Dependency Management를 참조하세요. `--packages` 옵션은 퍼블릭 Maven 리포지토리에 저장된 종속성을 지정합니다.

- `spark-redshift.jar`
- `spark-avro.jar`
- `RedshiftJDBC.jar`
- `minimal-json.jar`

Amazon EMR 릴리스 6.10.0 이상에서는 `minimal-json.jar` 종속성이 필요하지 않으며 기본적으로 다른 종속성을 각 클러스터에 자동으로 설치합니다. 다음 예제에서는 Apache Spark용 Amazon Redshift 통합을 사용하여 Spark 애플리케이션을 시작하는 방법을 보여줍니다.

Amazon EMR 6.10.0 +

다음 예제는 Amazon EMR 릴리스 6.10 이상에서 `spark-redshift` 커넥터를 사용하여 Spark 애플리케이션을 시작하는 방법을 보여줍니다.

```
spark-submit my_script.py
```

Amazon EMR 6.4.0 - 6.9.x

Amazon EMR 릴리스 6.4~6.9에서 `spark-redshift` 커넥터를 사용하여 Spark 애플리케이션을 시작하려면 다음 예제와 같이 `--jars` 또는 `--packages` 옵션을 사용해야 합니다. `--jars` 옵션과 함께 나열된 경로는 JAR 파일의 기본 경로입니다.

```
spark-submit \
  --jars /usr/share/aws/redshift/jdbc/RedshiftJDBC.jar,/usr/share/aws/redshift/
spark-redshift/lib/spark-redshift.jar,/usr/share/aws/redshift/spark-redshift/lib/
spark-avro.jar,/usr/share/aws/redshift/spark-redshift/lib/minimal-json.jar \
  my_script.py
```

Apache Spark용 Amazon Redshift 통합으로 인증

AWS Secrets Manager 를 사용하여 자격 증명을 검색하고 Amazon Redshift에 연결

다음 코드 샘플은 AWS Secrets Manager 를 사용하여 자격 증명을 검색하여 Python의 Apache Spark용 PySpark 인터페이스를 사용하여 Amazon Redshift 클러스터에 연결하는 방법을 보여줍니다.

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

secretsmanager_client = boto3.client('secretsmanager')
secret_manager_response = secretsmanager_client.get_secret_value(
    SecretId='string',
    VersionId='string',
```

```
    VersionStage='string'
)
username = # get username from secret_manager_response
password = # get password from secret_manager_response
url = "jdbc:redshift://redshifthost:5439/database?user=" + username + "&password=" +
 password

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3://path/for/temp/data") \
    .load()
```

IAM을 사용하여 보안 인증을 검색하고 Amazon Redshift에 연결

Amazon Redshift 제공 JDBC 드라이버 버전 2 드라이버를 사용하여 Spark 커넥터로 Amazon Redshift에 연결할 수 있습니다. AWS Identity and Access Management (IAM)를 사용하려면 IAM 인증을 사용하도록 JDBC URL을 구성합니다. Amazon EMR에서 Redshift 클러스터에 연결하려면 임시 IAM 보안 인증을 검색할 권한을 IAM 역할에 부여해야 합니다. 보안 인증을 검색하고 Amazon S3 작업을 실행할 수 있도록 IAM 역할에 다음 권한을 할당합니다.

- Redshift:GetClusterCredentials(프로비저닝된 Amazon Redshift 클러스터용)
- Redshift:DescribeClusters(프로비저닝된 Amazon Redshift 클러스터용)
- Redshift:GetWorkgroup(Amazon Redshift Serverless 작업 그룹용)
- Redshift:GetCredentials(Amazon Redshift Serverless 작업 그룹용)
- s3:GetBucket
- s3:GetBucketLocation
- s3:GetObject
- s3:PutObject
- s3:GetBucketLifecycleConfiguration

`GetClusterCredentials`에 대한 자세한 내용은 `GetClusterCredentials`에 대한 리소스 정책을 참조하세요.

또한 `COPY` 및 `UNLOAD` 작업 중에 Amazon Redshift가 IAM 역할을 맡을 수 있는지 확인해야 합니다.

JSON

```
{
    "Version": "2012-10-17",
    "Statement": [
        {
            "Effect": "Allow",
            "Principal": {
                "Service": "redshift.amazonaws.com"
            },
            "Action": "sts:AssumeRole"
        }
    ]
}
```

다음 예제에서는 Spark와 Amazon Redshift 사이에서 IAM 인증을 사용합니다.

```
from pyspark.sql import SQLContext
import boto3

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshift-host:redshift-port/db-name"
iam_role_arn = "arn:aws:iam::account-id:role/role-name"

# Read data from a table
df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("aws_iam_role", iam_role_arn) \
    .option("dbtable", "my_table") \
    .option("tempdir", "s3a://path/for/temp/data") \
    .mode("error") \
    .load()
```

Amazon Redshift에서 읽고 쓰기

다음 코드 예제는 데이터 소스 API와 SparkSQL을 통해 Amazon Redshift 데이터베이스에서 샘플 데이터를 읽고 쓰는 데 PySpark를 사용합니다.

Data source API

PySpark를 사용하여 데이터 소스 API를 통해 Amazon Redshift 데이터베이스에서 샘플 데이터를 읽고 씁니다.

```
import boto3
from pyspark.sql import SQLContext

sc = # existing SparkContext
sql_context = SQLContext(sc)

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

df = sql_context.read \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .load()

df.write \
    .format("io.github.spark_redshift_community.spark.redshift") \
    .option("url", url) \
    .option("dbtable", "tableName_copy") \
    .option("tempdir", "s3://path/for/temp/data") \
    .option("aws_iam_role", "aws_iam_role_arn") \
    .mode("error") \
    .save()
```

SparkSQL

PySpark를 사용하여 SparkSQL을 통해 Amazon Redshift 데이터베이스에서 샘플 데이터를 읽고 씁니다.

```
import boto3
import json
import sys
import os
from pyspark.sql import SparkSession

spark = SparkSession \
```

```
    .builder \
    .enableHiveSupport() \
    .getOrCreate()

url = "jdbc:redshift:iam://redshifthost:5439/database"
aws_iam_role_arn = "arn:aws:iam::accountID:role/roleName"

bucket = "s3://path/for/temp/data"
tableName = "tableName" # Redshift table name

s = f"""CREATE TABLE IF NOT EXISTS {tableName} (country string, data string)
    USING io.github.spark_redshift_community.spark.redshift
    OPTIONS (dbtable '{tableName}', tempdir '{bucket}', url '{url}', aws_iam_role
 '{aws_iam_role_arn}' ); """

spark.sql(s)

columns = ["country" ,"data"]
data = [("test-country","test-data")]
df = spark.sparkContext.parallelize(data).toDF(columns)

# Insert data into table
df.write.insertInto(tableName, overwrite=False)
df = spark.sql(f"SELECT * FROM {tableName}")
df.show()
```

Spark 커넥터 사용 시 고려 사항 및 제한 사항

- Amazon EMR의 Spark에서 Amazon Redshift로의 JDBC 연결을 위해 SSL을 켜는 것이 좋습니다.
- 모범 사례로 AWS Secrets Manager 에서 Amazon Redshift 클러스터의 보안 인증을 관리하는 것이 좋습니다. 예제AWS Secrets Manager 는를 사용하여 Amazon Redshift에 연결하기 위한 자격 증명 검색을 참조하세요.
- Amazon Redshift 인증 파라미터에 대해 `aws_iam_role` 파라미터를 사용하여 IAM 역할을 전달하는 것이 좋습니다.
- `tempdir` URI는 Amazon S3 위치를 가리킵니다. 이 임시 디렉터리는 자동으로 정리되지 않으므로, 추가 비용이 발생할 수 있습니다.
- Amazon Redshift에 대한 다음 권장 사항을 고려합니다.
  - Amazon Redshift 클러스터에 대한 퍼블릭 액세스를 차단하는 것이 좋습니다.
  - Amazon Redshift 감사 로깅을 켜는 것이 좋습니다.

  - Amazon Redshift 저장 데이터 암호화를 켜는 것이 좋습니다.
- Amazon S3에 대한 다음 권장 사항을 고려합니다.
  - Amazon S3 버킷에 대한 퍼블릭 액세스를 차단하는 것이 좋습니다.
  - Amazon S3 서버 측 암호화를 사용하여 사용된 Amazon S3 버킷을 암호화하는 것이 좋습니다.
  - Amazon S3 수명 주기 정책을 사용하여 Amazon S3 버킷에 대한 보존 규칙을 정의하는 것이 좋습니다.
  - Amazon EMR은 오픈 소스에서 이미지로 가져온 코드를 항상 확인합니다. 보안을 위해 Spark에서 Amazon S3로의 다음 인증 방법은 지원되지 않습니다.
    - `hadoop-env` 구성 분류에서 AWS 액세스 키 설정
    - `tempdir` URI에서 AWS 액세스 키 인코딩

커넥터 사용 및 지원되는 파라미터에 대한 자세한 내용은 다음 리소스를 참조하세요.

- Amazon Redshift 관리 안내서의 Apache Spark용 Amazon Redshift 통합
- Github의 `spark-redshift` community repository

Spark 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Spark의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

> **Important**
>
> Amazon EMR 릴리스 5.16.0부터 사용할 수 있는 Apache Spark 버전 2.3.1은 CVE-2018-8024 및 CVE-2018-1334를 처리합니다. Spark의 이전 버전을 Spark 버전 2.3.1 이상으로 마이그레이션하는 것이 좋습니다.

Spark 버전 정보

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-7.9.0	3.5.5	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.8.0	3.5.4	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.7.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
		datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.6.0	3.5.3	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-7.5.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.4.0	3.5.2	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-7.3.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.2.0	3.5.1	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.36.2	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-7.1.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-7.0.0	3.5.0	delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.15.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.14.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.13.0	3.4.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.12.0	3.4.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.11.1	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.11.0	3.3.2	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.10.1	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.10.0	3.3.1	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.9.1	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.9.0	3.3.0	aws-sagemaker-spark-sdk, delta, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.8.1	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.8.0	3.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.7.0	3.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.36.1	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.36.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.6.0	3.2.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.35.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.5.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, iceberg, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.4.0	3.1.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.3.1	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.3.0	3.1.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.2.1	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.2.0	3.0.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.1.1	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.1.0	3.0.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-6.0.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-6.0.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.34.0	2.4.8	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.33.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.33.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.32.1	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.32.0	2.4.7	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.31.1	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.31.0	2.4.6	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.2	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.30.1	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.30.0	2.4.5	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-notebook-env, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.29.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.28.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.28.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.27.1	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.27.0	2.4.4	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.26.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.25.0	2.4.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.24.1	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.24.0	2.4.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.23.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.23.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.22.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.21.2	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.21.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.21.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.20.1	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.20.0	2.4.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.19.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.19.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.18.1	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.18.0	2.3.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, nginx, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.2	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.17.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.17.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, emr-s3-select, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.16.1	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.16.0	2.3.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.15.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.15.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.14.2	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.14.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.14.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.13.1	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.13.0	2.3.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.12.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.12.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.12.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.4	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.11.3	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.2	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.11.1	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.11.0	2.2.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.10.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.10.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.9.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.9.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.8.3	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.2	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.8.1	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.8.0	2.2.0	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.7.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.7.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.6.1	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.6.0	2.1.1	emrfs, emr-goodies, emr-ddb, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.5.4	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.3	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.5.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.5.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.4.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.4.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.2	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.3.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.3.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.3	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.2	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.2.1	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.2.0	2.0.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.1.1	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.1.0	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.3	2.0.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.2	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-5.0.1	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-5.0.0	2.0.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.6	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.9.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.9.2	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.9.1	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.5	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.8.4	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.3	1.6.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.8.1	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.8.0	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.4	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.7.3	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.2	1.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.7.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.7.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.1	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.6.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.5.0	1.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.4.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.3.0	1.6.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-4.2.0	1.5.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.1.0	1.5.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave
emr-4.0.0	1.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave

Apache Sqoop

Note

Apache Sqoop을 포함하는 마지막 릴리스는 Amazon EMR 릴리스 7.4였습니다. Apache Sqoop은 Amazon EMR 릴리스 7.5 이상 릴리스에서 제거됩니다.

Apache Sqoop은 Amazon S3, Hadoop, HDFS 및 RDBMS 데이터베이스 간에 데이터를 전송하기 위한 도구입니다. 자세한 내용은 Apache Sqoop 웹 사이트를 참조하세요. Sqoop은 Amazon EMR 릴리스 5.0.0 이상에 포함되어 있습니다. 이전 릴리스에는 Sqoop이 샌드박스 애플리케이션으로 포함되었습니다. 자세한 내용은 Amazon EMR 4.x 릴리스 버전 단원을 참조하십시오.

주제

Sqoop 버전 정보

버전 7.4용 Sqoop 버전.

다음 테이블에는 Amazon EMR이 Sqoop를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Sqoop의 버전이 나열되어 있습니다.

emr-7.4.0용 Spark 버전 정보

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
emr-7.4.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodema

Amazon EMR 릴리스 레이블	Spark 버전	Spark와 함께 설치된 구성 요소
		nager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

6.15.0용 Sqoop 버전

다음 테이블에는 Amazon EMR이 Sqoop를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Sqoop의 버전이 나열되어 있습니다.

이 릴리스에서 Sqoop과 함께 설치된 구성 요소의 버전은 [릴리스 6.15.0 구성 요소 버전]을 참조하세요.

emr-6.15.0용 Sqoop 버전 정보

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.15.0	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

5.36.2용 Sqoop 버전

다음 테이블에는 Amazon EMR이 Sqoop를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Sqoop의 버전이 나열되어 있습니다.

이 릴리스에서 Sqoop과 함께 설치된 구성 요소의 버전은 [릴리스 5.36.2 구성 요소 버전]을 참조하세요.

emr-5.36.2용 Sqoop 버전 정보

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.36.2	Sqoop 1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR에서 Sqoop에 대한 고려 사항

Amazon EMR에서 Sqoop을 실행할 때는 다음 사항을 고려합니다.

HCatalog 통합을 통해 Sqoop 사용

Amazon EMR 기반 Sqoop은 Sqoop-HCatalog 통합을 지원합니다. Sqoop을 사용하여 Amazon S3의 HCatalog 테이블에 출력을 쓰는 경우 `mapred.output.direct.NativeS3FileSystem` 및 `mapred.output.direct.EmrFileSystem` 속성을 `false`로 설정하여 Amazon EMR 직접 쓰기를 비활성화합니다. 자세한 내용은 HCatalog 사용 단원을 참조하십시오. Hadoop `-D mapred.output.direct.NativeS3FileSystem=false` 및 `-D mapred.output.direct.EmrFileSystem=false` 명령을 사용할 수 있습니다. 직접 쓰기를 비활성화하지 않으면 오류는 발생하지 않지만 Amazon S3에 테이블이 생성되고 데이터는 기록되지 않습니다.

Sqoop JDBC 및 데이터베이스 지원

기본적으로 Sqoop에는 MariaDB 및 PostgreSQL 드라이버가 설치되어 있습니다. Sqoop용으로 설치된 PostgreSQL 드라이버는 PostgreSQL 8.4에서만 작동합니다. Sqoop용 대체 JDBC 커넥터 세트를 설치하려면 클러스터 마스터 노드에 연결하고 `/usr/lib/sqoop/lib`에 설치하십시오. 다음은 다양한 JDBC 커넥터에 대한 링크입니다.

- MariaDB: About MariaDB Connector/J를 참조하십시오.
- PostgreSQL: PostgreSQL JDBC 드라이버.
- SQLServer: SQL Server용 Microsoft JDBC Driver 다운로드.
- MySQL: Download Connector/J
- Oracle: Oracle Maven 리포지토리에서 Oracle JDBC 드라이버 및 UCP 가져오기

Sqoop용 지원 데이터베이스는 http://sqoop.apache.org/docs/*version*/SqoopUserGuide.html#_supported_databases에서 확인할 수 있습니다. *version*은 현재 사용하는 Sqoop의 버전입니다(예: 1.4.6). JDBC 연결 문자열이 이 목록에 나와 있는 것과 일치하지 않을 경우 드라이버를 지정해야 합니다.

예를 들면 다음 명령을 사용하여 Amazon Redshift 데이터베이스 테이블로 내보낼 수 있습니다(JDBC 4.1용).

```
sqoop export --connect jdbc:redshift://$MYREDSHIFTHOST:5439/mydb --
table mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 com.amazon.redshift.jdbc41.Driver --username master --password Mymasterpass1
```

MariaDB 및 MySQL 연결 문자열을 둘 다 사용할 수는 있지만 MariaDB 연결 문자열을 지정할 경우 해당 드라이버를 지정해야 합니다.

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

Secure Socket Layer 암호화를 사용하여 데이터베이스에 액세스하려면 다음과 같은 Sqoop 내보내기 예제에서처럼 JDBC URI를 사용해야 합니다.

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb?
verifyServerCertificate=false&useSSL=true&requireSSL=true --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver --
username master --password Mymasterpass1
```

RDS에서 SSL 암호화에 대한 자세한 내용은 Amazon RDS 사용 설명서에서 SSL을 사용하여 DB 인스턴스에 대한 연결 암호화를 참조하세요.

자세한 내용은 Apache Sqoop 설명서를 참조하십시오.

암호 변경

몇 가지 방법 중 하나를 선택하여 암호를 안전하게 전달할 수 있습니다.

Java KeyStore

선택한 방법에서는 JKS(Java KeyStore)를 사용하여 암호를 암호화하므로 암호를 읽을 수 있는 형식으로 저장할 필요가 없습니다.

1. 암호 별칭을 만듭니다. 메시지가 표시되면 데이터베이스에 액세스하는 데 사용하는 암호를 입력합니다.

```
hadoop credential create mydb.password.alias -provider jceks://hdfs/user/root/
mysql.password.jceks
```

2. 암호 별칭을 사용하여 Sqoop 작업을 시작합니다.

```
sqoop export -Dhadoop.security.credential.provider.path=jceks://hdfs/user/
root/mysql.password.jceks --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
 mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 org.mariadb.jdbc.Driver --username master --password-alias mydb.password.alias
```

--password-file

다음 예시에 보이는 것처럼 `--password-file` 명령을 사용하여 파일을 통해 암호를 전달할 수 있습니다.

1. 암호가 포함된 새로운 파일을 생성합니다.

```
echo -n 'Mymasterpass1' > /home/hadoop/mysql-pass.password
```

2. 파일을 사용하여 Sqoop 작업을 시작합니다.

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table
 mysqoopexport --export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver
 org.mariadb.jdbc.Driver --username master --password-file /home/hadoop/mysql-
pass.password
```

-P

다음 예시에 보이는 것처럼 -P 명령을 사용하여 프롬프트를 통해 암호를 입력할 수 있습니다.

```
sqoop export --connect jdbc:mariadb://$HOSTNAME:3306/mydb --table mysqoopexport --
export-dir s3://amzn-s3-demo-bucket/myinputfiles/ --driver org.mariadb.jdbc.Driver
 --username master -P
```

Sqoop 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Sqoop의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Sqoop 버전 정보

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-7.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.36.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-7.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-7.0.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.13.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.12.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.11.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.11.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.10.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.10.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.9.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.9.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.8.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.8.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.7.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.36.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.36.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.6.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.35.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.5.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.4.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.3.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.3.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.2.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.2.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-6.1.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-6.1.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.34.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.33.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.33.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.32.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.32.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.31.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.31.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.30.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.30.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.30.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mariadb-server, sqoop-client
emr-5.29.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.28.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.28.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.27.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.27.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.26.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.25.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.24.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.24.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.23.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.23.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.22.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.21.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.21.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.21.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.20.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.20.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.19.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.19.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.18.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.18.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.17.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.17.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.17.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.16.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.16.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.15.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.15.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.14.2	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.14.1	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.14.0	1.4.7	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.13.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.13.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.12.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.12.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.12.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.11.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.11.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.11.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.11.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.10.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.10.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.9.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.9.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.8.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.8.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.8.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.7.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.7.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.6.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.6.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, mysql-server, sqoop-client
emr-5.5.4	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.5.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.5.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.4.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.4.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.3.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.3.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.2.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.2.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.1.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.1.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.3	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

Amazon EMR 릴리스 레이블	Sqoop 버전	Sqoop과 함께 설치된 구성 요소
emr-5.0.2	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.1	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client
emr-5.0.0	1.4.6	emrfs, emr-ddb, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, mysql-server, sqoop-client

TensorFlow

TensorFlow는 머신 인텔리전스 및 딥 러닝 애플리케이션을 위한 오픈 소스 기호 수학 라이브러리입니다. 자세한 내용은 TensorFlow 웹 사이트를 참조하세요. TensorFlow는 Amazon EMR 릴리스 버전 5.17.0 이상에서 사용할 수 있습니다.

다음 테이블에는 Amazon EMR이 TensorFlow를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 TensorFlow의 버전이 나열되어 있습니다.

이 릴리스에서 TensorFlow와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 TensorFlow 버전 정보

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-7.9.0	TensorFlow 2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

다음 테이블에는 Amazon EMR이 TensorFlow를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 TensorFlow의 버전이 나열되어 있습니다.

이 릴리스에서 TensorFlow와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 TensorFlow 버전 정보

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.15.0	TensorFlow 2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

다음 테이블에는 Amazon EMR이 TensorFlow를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 TensorFlow의 버전이 나열되어 있습니다.

이 릴리스에서 TensorFlow와 함께 설치된 구성 요소의 버전은 [릴리스 5.36.2 구성 요소 버전](#)을 참조하세요.

emr-5.36.2용 TensorFlow 버전 정보

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.36.2	TensorFlow 2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EC2 인스턴스 유형별 TensorFlow 빌드

Amazon EMR은 클러스터에 대해 선택한 인스턴스 유형에 따라 다른 빌드의 TensorFlow 라이브러리를 사용합니다. Amazon EMR은 EMR-7.5.0 이상용 aarch64(Graviton) 인스턴스 유형이 있는 클러스터에 TensorFlow를 지원합니다. 다음 표에는 인스턴스 유형별 빌드가 나와 있습니다.

EC2 인스턴스 유형	TensorFlow 빌드
M5 및 C5	Intel MKL 최적화가 적용된 Tensorflow 2.16.1
P2, P4D, P5, G4DN, G5, G6, GR6	CUDA 12.3, cuDNN 8.9.7.29가 적용된 Tensorflow 2.16.1
P3, P3DN, G3, G3S	CUDA 12.3, cuDNN 8.9.7.29, NCCL 2.20.3-1이 적용된 Tensorflow 2.16.1 Nvidia NCCL은 P3 인스턴스에서만 사용할 수 있습니다. 최종 사용자 라이선스 계약(EULA): Amazon EMR에서 Nvidia 구성 요소를 사용하면 제품 EULA에 나와 있는 이용 약관에 동의하는 것입니다.
Graviton 인스턴스를 제외한 기타 모든 항목	Tensorflow 2.16.1

보안

Using TensorFlow securely의 지침을 따르고 프라이빗 서브넷에서 클러스터를 시작하여 신뢰할 수 있는 출처로 액세스를 제한하는 것이 좋습니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon VPC 옵션을 참조하세요.

TensorBoard 사용

TensorBoard는 TensorFlow 프로그램용 시각화 도구 제품군입니다. 자세한 내용은 Tensorflow 웹 사이트에서 TensorBoard: Visualized learning을 참조하세요.

TensorBoard를 Amazon EMR과 함께 사용하려면 클러스터 프라이머리 노드에서 TensorBoard를 시작해야 합니다.

Amazon EMR에서 Tensorboard를 Tensorflow와 함께 사용하는 방법

1. SSH를 사용하여 클러스터의 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. 다음 명령을 입력하여 마스터 노드에서 Tensorboard를 시작합니다. 요약 쓰기를 사용하여 요약 데이터를 생성하고 저장한 마스터 노드의 디렉터리로 */my/log/directory*를 바꿉니다.

   Amazon EMR 5.19.0 and later

   ```
   python3 -m tensorboard.main --logdir=/home/hadoop/tensor --bind_all
   ```

   Amazon EMR 5.18.1 and earlier

   ```
   python3 -m tensorboard.main --logdir=/my/log/dir
   ```

   기본적으로 마스터 노드는 포트 6006과 마스터 퍼블릭 DNS 이름을 사용하여 TensorBoard를 호스팅합니다. TensorBoard를 시작한 후 다음 예제와 같이 TensorBoard에 연결하는 데 사용할 수 있는 URL이 명령줄 출력에 표시됩니다.

   ```
   TensorBoard 2.16.1 at http://master-public-dns-name:6006 (Press CTRL+C to quit)
   ```
3. 신뢰할 수 있는 클라이언트의 마스터 노드에서 웹 인터페이스에 액세스하도록 설정합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon EMR 클러스터에 호스팅된 웹 인터페이스 보기를 참조하세요.
4. `http://master-public-dns-name:6006`에서 TensorBoard를 엽니다.

TensorFlow 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 TensorFlow의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

TensorFlow 버전 정보

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-7.9.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.8.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.7.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.6.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
		hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.5.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.4.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-7.3.0	2.16.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.2.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.2	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-7.1.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-7.0.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.15.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.14.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.13.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.12.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.11.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.11.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.10.1	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.10.0	2.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.1	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.9.0	2.10.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.8.1	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.8.0	2.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.7.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.36.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.36.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.6.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.35.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.5.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.4.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.3.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.3.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.2.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.2.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.1.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-6.0.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-6.0.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.34.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.33.1	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.33.0	2.4.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.32.1	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.32.0	2.3.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.1	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.31.0	2.1.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.30.2	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.30.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.29.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.28.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.27.1	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.27.0	1.14.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.26.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.25.0	1.13.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.24.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.23.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.23.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.22.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.21.2	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.21.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.20.1	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.20.0	1.12.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.19.1	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.19.0	1.11.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.18.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Amazon EMR 릴리스 레이블	TensorFlow 버전	TensorFlow와 함께 설치된 구성 요소
emr-5.17.2	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.1	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow
emr-5.17.0	1.9.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tensorflow

Apache Tez

Apache Tez는 데이터 처리 작업에 대한 복합 비순환 방향 그래프(DAG)를 생성하는 프레임워크입니다. 일부 사용 사례에서는 Hadoop MapReduce의 대안으로 사용할 수 있습니다. 예를 들면 Pig 및 Hive 워크플로는 Hadoop MapReduce를 사용하여 실행하거나 Tez를 실행 엔진으로 사용할 수 있습니다. 자세한 내용은 https://tez.apache.org/를 참조하십시오. Amazon EMR 릴리스 4.7.0 이상에는 Tez가 포함되어 있습니다.

다음 테이블에는 Amazon EMR이 Tez를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Tez의 버전이 나열되어 있습니다.

이 릴리스에서 Tez와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Tez 버전 정보

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-7.9.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

다음 테이블에는 Amazon EMR이 Tez를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Tez의 버전이 나열되어 있습니다.

이 릴리스에서 Tez와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Tez 버전 정보

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.15.0	Tez 0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
		hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

다음 테이블에는 Amazon EMR이 Tez를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Tez의 버전이 나열되어 있습니다.

이 릴리스에서 Tez와 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Tez 버전 정보

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.36.2	Tez 0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

주제

- Tez를 포함하는 클러스터 생성
- Tez 구성
- Tez 웹 UI
- Timeline Server
- Tez 릴리스 기록

Tez를 포함하는 클러스터 생성

Tez를 설치하려면 클러스터를 생성할 때 Apache Tez를 애플리케이션으로 선택합니다.

콘솔을 사용하여 Tez가 설치된 클러스터를 생성하려면

1. 새 Amazon EMR 콘솔로 이동하고 측면 탐색에서 이전 콘솔로 전환을 선택합니다. 이전 콘솔로 전환할 때 예상되는 사항에 대한 자세한 내용은 이전 콘솔 사용을 참조하세요.
2. 클러스터 생성 및 Go to advanced options(고급 옵션으로 이동)를 선택합니다.
3. 소프트웨어 구성에서 emr-4.7.0 이상의 릴리스를 선택합니다.
4. Amazon EMR에서 설치하려는 다른 애플리케이션과 함께 Tez를 선택합니다.
5. 필요에 따라 다른 옵션을 선택한 다음 클러스터 생성을 선택합니다.

를 사용하여 Tez로 클러스터를 생성하려면 AWS CLI

- `create-cluster` 명령을 `-- applications` 옵션과 함께 사용하여 Tez를 지정합니다. 다음은 Tez가 설치된 클러스터를 생성하는 예제입니다.

  > **Note**
  >
  > 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

  ```
  aws emr create-cluster --name "Cluster with Tez" --release-label emr-7.9.0 \
  --applications Name=Tez --ec2-attributes KeyName=myKey \
  --instance-type m5.xlarge --instance-count 3 --use-default-roles
  ```

Tez 구성

`tez-site` 구성 분류를 통해 값을 설정하여 Tez를 사용자 지정할 수 있습니다. 이는 `tez-site.xml` 구성 파일의 설정을 구성합니다. 자세한 내용은 Apache Tez 설명서의 TezConfiguration을 참조하십시오. Tez 실행 엔진을 사용하도록 Hive 또는 Pig를 변경하려면 `hive-site` 및 `pig-properties` 구성 분류를 적절하게 사용합니다. 예제는 아래에 나와 있습니다.

구성의 예

Example 예제: Tez 루트 로깅 수준 사용자 지정과 Hive 및 Pig에 대한 실행 엔진으로 Tez 설정

아래에 표시된 예제 `create-cluster` 명령은 Tez, Hive 및 Pig가 설치된 클러스터를 생성합니다. 이 명령은 Amazon S3에 저장된 파일(`myConfig.json`)을 참조합니다. 이 파일은 `hive-site` 및 `pig-properties` 구성 분류를 사용하여 실행 엔진을 Pig 및 Hive에 대해 Tez를 설정하고 `tez.am.log.level`을 `DEBUG`로 설정하는 `tez-site` 분류의 속성을 지정합니다.

> **Note**
>
> 가독성을 위해 Linux 줄 연속 문자(\)가 포함됩니다. Linux 명령에 사용하거나 제외할 수 있습니다. Windows에서는 제외시키거나 캐럿(^)으로 바꿉니다.

```
aws emr create-cluster --release-label emr-7.9.0 \
--applications Name=Tez Name=Hive Name=Pig --ec2-attributes KeyName=myKey \
--instance-type m5.xlarge --instance-count 3 \
--configurations https://s3.amazonaws.com/amzn-s3-demo-bucket/myfolder/myConfig.json --
use-default-roles
```

다음은 `myConfig.json`의 콘텐츠 예제입니다.

```
[
    {
      "Classification": "tez-site",
      "Properties": {
        "tez.am.log.level": "DEBUG"
      }
    },
    {
      "Classification": "hive-site",
      "Properties": {
        "hive.execution.engine": "tez"
      }
    },
    {
      "Classification": "pig-properties",
      "Properties": {
        "exectype": "tez"
```

```
      }
    }
  ]
```

> **Note**
>
> Amazon EMR 버전 5.21.0 이상에서는 클러스터 구성을 재정의할 수 있으며, 실행 중인 클러스터의 각 인스턴스 그룹에 대해 추가 구성 분류를 지정할 수 있습니다. Amazon EMR 콘솔, AWS Command Line Interface (AWS CLI) 또는 AWS SDK를 사용하여이 작업을 수행할 수 있습니다. 자세한 내용은 실행 중 클러스터의 인스턴스 그룹에 대해 구성 제공을 참조하십시오.

Tez 비동기 분할 열기

테이블 경로에 많은 수의 작은 파일이 있고 쿼리가 이를 모두 읽으려고 시도하는 경우 각 개별 분할에 대응되는 각 작은 파일은 하나의 Tez 그룹화된 분할 아래에 결합됩니다. 그 다음에는 단일 매퍼가 단일 Tez 그룹화된 분할을 처리합니다. 실행은 동기식이기 때문에 그룹화된 분할 아래의 각 개별 분할은 하나씩 처리됩니다. 이 작업을 하려면 `RecordReader` 객체가 분할을 동기적으로 처리해야 합니다.

Amazon EMR 6.15.0에는 Tez 그룹화된 분할에서 입력 분할을 비동기적으로 열도록 지정할 수 있는 구성이 도입되었습니다. 이 특성은 TEZ-4397에 의해 시작되었지만 OSS Hive에서는 회귀가 발생했습니다. EMR Hive는 Hive ACID 테이블에서 회귀 및 추가 버그를 수정했습니다. 이렇게 개선하면 단일 Tez 그룹화된 분할에 많은 수의 입력 분할이 있을 때 읽기 쿼리 성능이 빨라집니다.

명칭	분류	설명
`tez.grouping.split.init.threads`	`tez-site`	Tez가 `RecordReaders` 를 사전 개시하고 분할을 여는 데 사용하는 대몬(daemon) 스레드 수를 지정합니다. ACID 테이블의 경우 `tez.grouping.split.init.threads` 의 지원되는 최대값은 1입니다.
`tez.grouping.split.init.recordreaders`	`tez-site`	대몬(daemon) 스레드에서 사전 초기화한 상태로 유지할 `RecordReaders` 의 개수를

명칭	분류	설명
		지정합니다. 이는 Tez 그룹화된 분할에 `InputSplits` 가 다량 포함되어 있을 때 도움이 될 수 있습니다. 해당 입력 `RecordReaders` 분할을 처리하기 위한 초기화는 순차적 처리 대신 대몬(daemon) 스레드를 사용하여 비동기적으로 수행할 수 있습니다.

구성 참고 사항:

주제	세부 사항	
권장 구성 설정	위의 구성 설정을 `hive-site` 및 모두에서 원하는 값으로 설정하는 것이 좋습니다`tez-site`.	
일치하는 값	구성 설정의 값은 `hive-site` 및 모두에서 동일해야 합니다`tez-site`.	
LLAP 권장 사항	LLAP가 활성화된 경우이 기능을 사용하지 않는 것이 좋습니다.	

Tez 비동기 분할 열기를 위한 벤치마킹

Tez 비동기 분할 열기 기능을 벤치마킹하기 위해 다음 환경 및 구성을 사용했습니다.

- 벤치마크 환경 – m5.16xlarge를 사용하는 프라이머리 노드 1개와 m5.16xlarge를 사용하는 코어 노드 16개가 있는 Amazon EMR 클러스터입니다.
- 벤치마크 구성 – 단일 Tez 그룹화된 분할에 입력 분할이 많은 벤치마킹 시나리오를 시뮬레이션하기 위해 `tez.grouping.split-count`를 1로 설정합니다.

- 벤치마킹을 위해 사용되는 표 - 테이블에는 파티션 200개가 포함되어 있고 각 파티션에는 단일 파일이 포함되어 있습니다. 벤치마크는 해당 테이블에 CSV 파일이 포함된 경우와 해당 테이블에 파켓 파일이 포함된 경우를 위해 수행됩니다. 벤치마킹을 위한 Hive 쿼리: 테이블의 `SELECT COUNT(*)`를 10회 수행하고 평균 런타임을 계산합니다.
- Tez 비동기 분할 열기를 활성화하기 위한 구성 - 다음과 같습니다.
  - `tez.grouping.split.init.threads` = 4
  - `tez.grouping.split.init.recordreaders` = 10

데이터세트	기능 비활성화(기준)	기능 활성화	개선
CSV 데이터 세트	90.26초	79.20초	12.25%
Parquet 데이터 세트	54.67초	42.23초	22.75%

Tez 웹 UI

Tez에는 고유한 웹 사용자 인터페이스가 있습니다. 웹 UI를 보려면 다음 URL을 참조하세요.

```
http://masterDNS:8080/tez-ui
```

Tez 웹 UI에서 Hive Queries 탭을 활성화하려면 다음 구성을 설정합니다.

```
[
  {
    "Classification": "hive-site",
    "Properties": {
      "hive.exec.pre.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.post.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook",
      "hive.exec.failure.hooks": "org.apache.hadoop.hive.ql.hooks.ATSHook"
    }
  }
]
```

콘솔의 클러스터 세부 정보 페이지에서 애플리케이션 사용자 인터페이스 탭을 사용하여 Tez, Spark 및 YARN 애플리케이션 UI 세부 정보를 볼 수 있습니다. Amazon EMR 애플리케이션 사용자 인터페이스(UI)는 클러스터 외부에서 호스팅되며 클러스터가 종료된 후에도 사용할 수 있습니다. SSH 연결이나

웹 프록시를 설정할 필요가 없으므로 사용하면 활성 작업 및 작업 기록을 분석하여 보다 쉽게 문제를 해결할 수 있습니다.

자세한 내용은 Amazon EMR 관리 안내서에서 애플리케이션 기록 보기를 참조하세요.

Timeline Server

Tez가 설치될 때 실행하도록 YARN 타임라인 서버가 구성됩니다. 타임라인 서버를 사용하여 Tez나 MapReduce 실행 엔진을 통해 제출한 작업을 보려면 `http://master-public-DNS:8188` URL을 사용하여 웹 UI를 살펴봅니다. 자세한 내용은 Amazon EMR 관리 안내서에서 Amazon EMR 클러스터에 호스팅된 웹 인터페이스 보기를 참조하세요.

Tez 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Tez의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Tez 버전 정보

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-7.9.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.8.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
		server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.7.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.6.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-7.5.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.4.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-7.3.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-7.2.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-5.36.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-7.1.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-7.0.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.15.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.14.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.13.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.12.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.11.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.11.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker
emr-6.10.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn, tez-on-worker

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.9.1	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.9.0	0.10.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.8.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.8.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.7.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.36.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.36.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.6.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.35.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.5.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.4.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.3.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.3.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.2.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.1.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.1.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-6.0.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-6.0.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.34.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.33.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.33.0	0.9.2	emrfs, emr-goodies, hadoop-cl ient, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.1	0.9.2	emrfs, emr-goodies, hadoop-cl ient, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.32.0	0.9.2	emrfs, emr-goodies, hadoop-cl ient, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.31.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.31.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.2	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.30.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.30.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.29.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.28.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.28.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.27.1	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.27.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.26.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.25.0	0.9.2	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.24.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.24.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.23.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.23.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.22.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.2	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.21.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.21.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.20.1	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.20.0	0.9.1	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.19.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.18.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.18.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.17.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.17.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.16.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.16.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.15.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.14.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.14.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.13.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.13.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.12.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.12.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.11.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.11.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.11.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.10.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.10.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.9.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.7.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.6.1	0.8.4	emrfs, emr-goodies, hadoop-cl ient, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.6.0	0.8.4	emrfs, emr-goodies, hadoop-cl ient, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.4	0.8.4	emrfs, emr-goodies, hadoop-cl ient, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodema nager, hadoop-yarn-resour cemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.5.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.5.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.5.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.4.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.3.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.3.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.2.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.2.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.2.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.1.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.0.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-5.0.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-5.0.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.6	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-4.9.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.9.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-4.9.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.5	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.4	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-4.8.3	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.2	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.8.1	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-4.8.0	0.8.4	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.4	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.3	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

Amazon EMR 릴리스 레이블	Tez 버전	Tez와 함께 설치된 구성 요소
emr-4.7.2	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.1	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn
emr-4.7.0	0.8.3	emrfs, emr-goodies, hadoop-client, hadoop-mapred, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, tez-on-yarn

버전별 Tez 릴리스 정보

주제

- Amazon EMR 7.8.0 - Tez 릴리스 정보

Amazon EMR 7.8.0 - Tez 릴리스 정보

Amazon EMR 7.8.0 - Tez 변경 사항

유형	설명
개선 사항	하둡과 일치하도록 bouncycastle 버전을 1.78로 업그레이드했습니다.

Amazon EMR 7.7.0 - Tez 릴리스 정보

Amazon EMR 7.7.0 - Tez 변경 사항

유형	설명
개선 사항	Apache Tez의 `tez.task.relaxed.locality` 속성은 작업 예약이 데이터 로캘리티 제약 조건(랙 및 노드 로캘)을 엄격하게 따르는지 여부를 제어합니다. `true` (EMR-7.6+의 기본값)로 설정하면 Tez는 로캘리티를 적용하지 않으므로 사용 가능한 컨테이너에 태스크를 할당할 수 있으므로 리소스 사용률이 향상되고 사용 중인 클러스터의 대기 시간이 줄어듭니다.

Amazon EMR 7.6.0 - Tez 릴리스 정보

Amazon EMR 7.6.0 - Tez 변경 사항

유형	설명
개선 사항	의 기본값을 false`tez.task.relaxed.locality` 로 변경하여 작업용 컨테이너를 요청하는 동안에는 랙 및 노드 로캘리티 제약 조건을 고려하지 않습니다.
개선	로컬성으로 인한 지연을 비활성화하고 로컬이 아닌 폴백을 허용하도록 구성 조정
개선	TEZ-4547: JobConf JobID 추가

Amazon EMR 7.5.0 - Tez 릴리스 정보

Amazon EMR 7.5.0 - Tez 변경 사항

유형	설명
개선 사항	기본적으로 작업 선점 활성화
개선	비동기식으로 선제적으로 10개로 시작하도록 레코드 리더 수 증가
개선	TEZ-4580: 재사용이 활성화된 경우 새 컨테이너의 느린 선점

Amazon EMR 7.4.0 - Tez 릴리스 정보

Amazon EMR 7.4.0 - Tez 변경 사항

유형	설명
개선 사항	HADOOP-175244를 통해 Hadoop 3.4.0에서 제거된 Tez에서 EventCounter appender 제거
업그레이드	commons-cli를 1.5.0으로 업그레이드
업그레이드	bouncy-castle 버전을 1.70.0으로 업그레이드

Amazon EMR 7.2.0 - Tez 릴리스 정보

Amazon EMR 7.2.0 - Tez 변경 사항

유형	설명
개선 사항	작업용 컨테이너를 요청할 때 로캘리티 제약 조건을 비활성화하는 옵션이 제공됩니다.
버그 수정	TaskDependencies#addTaskDependency 스레드를 안전하게 설정

유형	설명
업그레이드	Tez용 TLS 버전을 1.3으로 업그레이드합니다.

Amazon EMR 7.2.0 - 새 구성

분류	명칭	Default	설명
tez-site	tez.task.relaxed.locality	false	활성화하면 태스크에 대한 컨테이너를 요청하는 동안 랙 및 노드 로캘 제약 조건이 고려되지 않습니다.

Amazon EMR 7.1.0 - Tez 릴리스 정보

Amazon EMR 7.1.0 - Tez 변경 사항

유형	설명
버그 수정	TEZ-4394 – Netty4 ShuffleHandler: 보스 스레드 1개를 사용해야 함(#256)
업그레이드	Netty를 4.1.100.Final로 업그레이드합니다.
업그레이드	Jetty를 9.4.53.v20231009로 업그레이드합니다.

Amazon EMR 7.0.0 - Tez 릴리스 정보

Amazon EMR 7.0.0 Tez에는 Amazon EMR-6.15.0 Tez를 포함하여 이루어진 모든 변경 사항 및 업데이트가 포함되어 있습니다.

Amazon EMR 7.0.0 - Tez 변경 사항

유형	설명
업그레이드	Tez 런타임은 이제 기본적으로 Java 17을 사용합니다. 자세한 내용은 EMR 7.0.0 릴리스 가이드를 참조하세요.

Amazon EMR 6.15.0 - Tez 릴리스 정보

Amazon EMR 6.15.0 - Tez 변경

유형	설명
Feature	TEZ-4397: Tez 입력 분할을 비동기적으로 엽니다.
업그레이드	TEZ-4493: Apache Hadoop을 3.3.6으로 업그레이드합니다.

Amazon EMR 6.15.0 - 테즈 기능

- Tez 비동기 분할 열기 – Amazon EMR 6.15.0에는 Tez 그룹화된 분할에서 입력 분할을 비동기식으로 열도록 지정할 수 있는 구성이 도입되었습니다. 이 특성은 TEZ-4397에 의해 시작되었지만 OSS Hive에서는 회귀가 발생했습니다. Amazon EMR Hive는 Hive ACID 테이블의 회귀 현상 및 추가 버그 문제를 해결했습니다. 이 개선 사항으로 인해 단일 Tez 그룹화된 분할에 많은 수의 입력 분할이 있을 경우 읽기 쿼리 성능이 더욱 빨라집니다. 자세한 내용은 Tez 비동기 분할 열기 단원을 참조하십시오.

Amazon EMR 6.14.0 - Tez 릴리스 정보

Amazon EMR 6.14.0 - Tez 변경

유형	설명
개선 사항	Tez의 TLS 버전을 1.2로 업그레이드

Amazon EMR 6.13.0 - Tez 릴리스 정보

Amazon EMR 6.13.0 - Tez 변경

유형	설명
버그 수정	TEZ-4295 되돌리기: 데이터 압축을 풀 수 없습니다. 버퍼 길이가 너무 작습니다.
버그 수정	TEZ-4302 되돌리기: `GzipCodec` 를 포함하는 `CodecUtils` 에서 `NullPointerException` 예외가 발생합니다.
버그 수정	TEZ-4234 되돌리기: 한도가 용량을 초과하는 컴프레서의 `Buffer.limit` 에서 `IllegalArgumentException` 예외가 발생할 수 있습니다.
버그 수정	TEZ-4135 되돌리기: 메모리 내 읽기 실행 시 메모리 할당을 개선합니다.

Amazon EMR 6.12.0 - Tez 릴리스 정보

Amazon EMR 6.12.0 - Tez 변경

유형	설명
개선 사항	JDK 11 및 JDK 17 런타임에 대한 지원 추가됨
버그 수정	TEZ-4492: Bower Registry CERT_EXPIRE 문제(BOWER-2608)를 방지하기 위해 bower.herokuapp 미러를 사용하도록 Bowerrc를 업데이트합니다.
업그레이드	Surefire를 3.0.0-M7로 업그레이드함

Amazon EMR 6.11.0 - Tez 릴리스 정보

Amazon EMR 6.11.0 - Tez 변경

유형	설명
버그	셔플 데이터의 버텍스 수준 정리 작업 중에 잘못된 버텍스 상태 전환을 수정함
버그	셔플 데이터의 DAG 또는 버텍스 수준 정리가 작동하지 않던 문제를 수정함
개선 사항	완료된 DAG의 셔플 데이터를 지우기 위해 기본적으로 tez.am.dag.cleanup.on.completion 활성화

Amazon EMR 6.10.0 - Tez 릴리스 정보

Amazon EMR 6.10.0 - Tez 변경

유형	설명
Feature	기본적으로 `tez.runtime.transfer.data-via-events.enabled` 활성화
백포트	TEZ-4450: 셔플 데이터가 데이터 이동 이벤트를 통해 전송될 때 나타나는 셔플 데이터 가져오기 실패 수정
백포트	TEZ-4460: Tez 셔플 핸들러에서 셔플 데이터를 가져올 때 발생하는 읽기 제한 시간 오류 수정
백포트	TEZ-4455: 더 나은 디버깅을 위해 ShuffleHandler 파이프라인에 LoggingHandler 추가
버그	태스크 선점이 활성화된 경우 Tez 작업이 간헐적으로 멈추는 문제 수정

Amazon EMR 6.9.0 - Tez 릴리스 정보

Amazon EMR 6.9.0 - Tez 변경

유형	설명
업그레이드	Tez를 0.10.2로 업그레이드했습니다. 자세한 내용은 change log for Apache Tez 0.10.2를 참조하세요.
업그레이드	Hadoop을 3.3.3으로 업그레이드합니다.
버그	TEZ-4450 때문에 기본적으로 `tez.runtime.transfer.data-via-events.enabled` 를 비활성화합니다.

Amazon EMR 6.8.0 - Tez 릴리스 정보

Amazon EMR 6.8.0 - Tez 변경

유형	설명
백포트	TEZ-3363: 셔플 핸들러의 버텍스 수준에서 중간 데이터 삭제
백포트	TEZ-4129: 셔플 핸들러의 실패한 시도에 대한 중간 시도 데이터 삭제
백포트	TEZ-4430: tez.task.launch.cmd-opts 속성이 작동하지 않는 문제 수정

Amazon EMR 6.7.0 - Tez 릴리스 정보

Amazon EMR 6.7.0 - Tez 변경

유형	설명
백포트	TEZ-4403: SLF4J 버전을 1.7.36으로 업그레이드
백포트	TEZ-4405: log4j 1.x를 reload4j로 바꾸기
백포트	TEZ-4411: Tez 빌드 실패: FileSaver.js를 찾을 수 없음

Amazon EMR 6.6.0 - Tez 릴리스 정보

Amazon EMR 6.6.0 - Tez 변경

유형	설명
백포트	TEZ-3918: tez.task.log.level 속성이 작동하지 않는 문제를 수정했습니다.
백포트	TEZ-4353: commons-io를 2.8.0로 업데이트합니다.
백포트	TEZ-4114: tez에서 직접적인 jetty 종속 항목을 제거합니다.
백포트	TEZ-4323: TEZ-4114의 dist 패키지에서 Jetty jar이 제거되었습니다.

Trino

Trino는 다양한 데이터 소스에 대한 대화형 쿼리를 위해 설계된 오픈 소스 쿼리 엔진입니다. 여기에는 관계형 데이터베이스, 파일 기반 데이터, HDFS 데이터 등이 포함될 수 있습니다. Amazon EMR을 사용하는 Trino의 가장 일반적인 목적은 Amazon S3에 저장된 대규모 데이터 세트에서 복잡한 SQL 쿼리를 실행하는 것입니다. 또한 ANSI SQL을 준수하므로 SQL에 익숙한 데이터베이스 엔지니어, 데이터 분석가 및 데이터 과학자에게 친숙합니다.

> **Note**
>
> PrestoSQL은 2020년 12월에 Trino로 이름이 바뀌었습니다. Amazon EMR 버전 6.4.0 이상은 일반적으로 Trino를 참조하는 반면, 이전 릴리스 버전은 PrestoSQL을 참조합니다.

> **Important**
>
> 이전 버전의 Trino인 PrestoSQL은 Amazon EMR에서 계속 사용할 수 있습니다. 하지만 앞으로 Amazon EMR에서 사용할 Trino를 적극 권장합니다. 또한 Trino와 PrestoSQL은 동일한 클러스터에서 동시에 실행할 수 없습니다.

다음 표에는 Amazon EMR이 Trino와 함께 설치하는 구성 요소와 함께 Amazon EMR 7.x의 최신 릴리스에 포함된 Trino의 버전이 나열되어 있습니다. 이 릴리스에서 Trino와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Trino 버전 정보

Amazon EMR 릴리스 레이블	Trino 버전	Trino와 함께 설치된 구성 요소
emr-7.9.0	Trino 467	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hive-client,

Amazon EMR 릴리스 레이블	Trino 버전	Trino와 함께 설치된 구성 요소
		hudi, hudi-trino, hcatalog-server, mariadb-server, trino-coordinator, trino-worker

주제

Trino 기록 및 설계

Trino는 다양한 소스에서 대규모 데이터 세트를 쿼리하는 데 특화되어 있습니다. Trino는 기존 빅 데이터 사용 사례에서 HDFS에 액세스하고 쿼리할 수 있지만 관계형 데이터베이스 및 NoSQL 데이터베이스와 같은 추가 소스를 쿼리할 수도 있습니다. Trino는 원래 2019년에 Presto 쿼리 엔진의 포크로 시작되었습니다. 이후 Presto 코드 베이스와 독립적으로 개발되었습니다.

Trino 쿼리 엔진 및 사용 방법에 대한 자세한 내용은 Trino 웹 사이트를 참조하세요. Trino 소스 설명서를 읽으려면 Trino 개요를 참조하세요.

아키텍처 개념

Trino는 클러스터 전체에서 데이터를 병렬로 처리하기 때문에 빠르고 효율적인 쿼리를 실행할 수 있습니다. 일반적으로 하둡 및 HDFS와 관련된 사용 사례에서 대용량 데이터 볼륨에 대한 쿼리를 전문으로 하기 때문에 데이터 레이크 쿼리를 염두에 두고 설계되었습니다. 그러나 기존 관계형 데이터베이스도 쿼리할 수 있습니다. 자세한 내용은 Trino 설명서의 아키텍처를 참조하세요.

Trino의 구성 요소

Trino에는 쿼리를 빠르게 실행하기 위해 함께 작동하는 몇 가지 주요 아키텍처 구성 요소가 있습니다. 더 나은 성능을 위해 클러스터를 미세 조정할 때 이에 대한 실무 지식을 갖추는 데 도움이 됩니다.

- 코디네이터는 쿼리 오케스트레이션을 담당합니다. 들어오는 SQL 쿼리를 구문 분석 및 최적화하고, 실행 계획을 생성하고, 작업자 노드에 작업을 할당하고, 쿼리 결과를 수집 및 수집합니다. 또한 리소스 사용량을 모니터링하고 작업자 노드의 상태를 추적합니다. 자세한 내용은 Trino 설명서의 Coordinator를 참조하세요.
- 작업자 노드는 쿼리에 대한 데이터 처리를 처리합니다. 코디네이터가 작업을 할당한 후 작업자는 데이터를 검색하고, 조인 및 집계와 같은 필요한 작업을 수행하고, 중간 데이터를 다른 작업자와 교환합니다. 자세한 내용은 Trino 설명서의 작업자를 참조하세요.
- 커넥터는 Trino가 다양한 데이터 소스에 연결하고 쿼리할 수 있는 플러그인입니다. 각 커넥터는 Amazon S3, Apache Hive 또는 관계형 데이터베이스와 같은 소스에서 데이터에 액세스하고 검색하는 방법을 알고 있습니다. 이러한 커넥터는 소스 데이터를 Trino의 스키마 구조에 매핑합니다.
- 카탈로그는 특정 커넥터와 연결된 스키마 및 테이블의 논리적 모음입니다. 조정자에 정의된 카탈로그를 사용하면 Trino가 다양한 데이터 소스를 단일 네임스페이스로 처리할 수 있습니다. 따라서 사용자는 동일한 쿼리에서 통합된 방식으로 Hive 및 MySQL과 같은 여러 소스를 함께 쿼리할 수 있습니다.
- Trino CLI와 같은 클라이언트는 JDBC 및 ODBC 드라이버를 통해 Trino 조정자에 연결하여 SQL 쿼리를 제출합니다. 코디네이터는 쿼리 수명 주기를 관리하여 추가 분석 또는 보고를 위해 클라이언트에 결과를 제공합니다.

쿼리 실행

Trino가 SQL 문을 가져와 쿼리로 실행하는 방법을 이해하려면 Trino 설명서의 Trino 개념을 참조하세요.

Trino 시작하기

이 섹션의 절차에서는 Trino를 사용하여 메타스토어 데이터 소스를 쿼리하기 위해 Amazon EMR 클러스터를 설정하는 방법을 보여줍니다. Glue 데이터 카탈로그를 포함하는 이러한 AWS 메타스토어는 메타데이터 및 데이터베이스 객체를 저장하고 액세스 권한을 관리합니다. 이 절차에서는 사전 조건, 권장 구성 설정, 커넥터 생성 및 메타스토어 테이블에서 쿼리 실행을 다룹니다.

주제

에서 Amazon EMR을 사용하기 위한 사전 조건 단계 완료 Trino

를 사용하지 AWS않았거나 Amazon EMR 클러스터를 생성하지 않은 경우 Trino를 사용하여 Amazon EMR 클러스터를 생성하기 전에 다음 사전 조건 단계를 완료하세요.

AWS 환경 설정

아직 AWS 계정을 구성하지 않았다면 다음 단계를 완료하세요.

1. AWS 아직 계정이 없는 경우 계정에 가입합니다. 자세한 내용은 AWS 계정 관리 참조 안내서의 계정 생성을 AWS 참조하세요.
2. 관리 사용자로 계정에 로그인합니다.
3. 그룹을 생성하고 그룹에 사용자를 할당합니다.
4. 나중에 SSH와의 리소스 간 통신을 보호하는 데 사용할 수 있는 Amazon EC2 키 페어를 생성합니다. 이 단계는 작업을 수행하기 위해 프라이머리 노드에 연결하려는 경우에 필요합니다. 자세한 내용은 SSH를 사용하여 Amazon EMR 클러스터 프라이머리 노드에 연결을 참조하세요.

Trino를 사용하여 Amazon EMR 클러스터 시작

다음은 Trino로 클러스터를 생성할 때 올바른 구성 선택에 대해 설명합니다.

Hive 커넥터를 사용하여 쿼리에 데이터를 사용할 수 있도록 설정

클러스터에서 메타스토어 데이터를 쿼리할 목적으로 Hive 메타스토어에 대한 Trino 커넥터를 구성할 수 있습니다. 메타스토어는 파일 기반 콘텐츠 또는 데이터를 테이블로 사용할 수 있도록 하는 추상화 계층이므로 쿼리하기 쉽습니다. 클러스터에서 Hive 메타스토어 테이블을 사용할 수 있도록 Amazon EMR에서 커넥터를 구성해야 합니다. 다음 절차에서는이 작업을 수행하는 방법을 보여줍니다.

1. 콘솔에서 AWS Glue를 선택하고 Amazon S3의 소스 데이터를 기반으로 테이블을 생성합니다. AWS Glue 데이터 카탈로그의 테이블은 데이터에 대한 메타데이터 정의입니다. 이 컨텍스트에서는 테이블을 수동으로 생성하여 소스 데이터에서 원하는 대로 열을 생성하는 것이 좋습니다. Amazon S3의 반정형 데이터에서 AWS Glue의 테이블을 생성하는 방법에 대한 자세한 내용은 AWS Glue 사용 설명서의 콘솔을 사용하여 테이블 생성을 참조하세요.
2. 클러스터 생성의 일부로 구성을 설정합니다. 구성 탭을 선택합니다. 구성은 클러스터의 선택적 사양입니다. 구성을 입력할 때 다음 샘플과 같이 JSON을 추가합니다.이 샘플은 Trino에게 Glue 데이터 카탈로그를 테이블 메타데이터에 대한 외부 Hive AWS 메타스토어로 사용하도록 지시합니다.

```
{
```

```
        "classification": "trino-connector-hive",
        "properties": {
            "hive.metastore": "glue"
        }
    }
```

또는 클러스터를 생성할 때 소프트웨어 설정 섹션에서 구성을 적용할 수 있습니다.

또한 Apache Iceberg 연결과 같은 다른 커넥터 유형을 설정할 수 있습니다. 자세한 내용은 Amazon EMR 릴리스 안내서의 Trino에서 Iceberg 클러스터 사용을 참조하세요. 추가 설정 구성은 선택 사항입니다.

시작하기 단계를 계속하려면 섹션을 참조하세요Amazon EMR 클러스터의 프라이머리 노드에 연결하고 쿼리를 실행합니다..

Trino를 사용하여 클러스터 생성

다음은 Trino와 함께 사용할 클러스터를 생성할 때 올바른 구성 선택에 대해 설명합니다.

> **Important**
>
> 클러스터를 생성하기 전에 Glue 데이터 카탈로그 구성을 Hive AWS 메타스토어로 완료합니다. 이 구성을 시작하는 것이 좋습니다. 자세한 내용은 Hive 커넥터를 사용하여 쿼리에 데이터를 사용할 수 있도록 설정 단원을 참조하십시오.

1. AWS 콘솔의 서비스에서 Amazon EMR을 선택합니다. Amazon EMR을 선택하면 기존 클러스터가 있는 경우 EC2 클러스터의 EMR이 나열됩니다.
2. 클러스터 생성을 선택합니다. 여기에서 클러스터 빌드 프로세스를 시작합니다.
3. 클러스터에 이름을 지정하고 Amazon EMR 릴리스를 선택합니다. 자습서의 최신 릴리스를 선택할 수 있습니다.
4. Trino 애플리케이션이 미리 선택된 Trino 번들을 선택합니다. 번들은 클러스터의 목적을 미리 알고 있을 때 편의를 위해 설정됩니다. 그렇지 않으면 Trino의 확인란을 선택하면 됩니다.
5. 클러스터 구성에서 균일한 인스턴스 그룹을 선택합니다. 계속해서 추가 인스턴스 그룹을 제거합니다.
6. 인스턴스 유형을 선택합니다. 일반적으로 메모리가 16GiB 이상인 인스턴스 유형을 선택하는 것이 좋습니다. 또한 클러스터 조정 및 프로비저닝에서 클러스터 크기 수동 설정을 선택합니다.

7. 이때 Glue를 가리키도록 Hive AWS 메타스토어 구성을 설정합니다. 자세한 내용은 단원을 참조하십시오[Hive 커넥터를 사용하여 쿼리에 데이터를 사용할 수 있도록 설정](). 클러스터를 빌드하기 전에 이 작업을 완료합니다.
8. 클러스터 생성을 선택합니다. 완료하는 데 몇 분 정도 걸릴 수 있습니다.

   이 단계에서는 모든 구성 단계를 자세히 다루지는 않습니다. 클러스터 설정에 대한 자세한 내용은 [Amazon EMR 클러스터 계획, 구성 및 시작에서 확인할 수 있습니다]().

> **Note**
>
> 동일한 클러스터에서 사용할 Presto와 Trino를 모두 선택하지 마십시오. 이들을 함께 실행하는 것은 지원되지 않습니다. Trino를 실행하는 경우 Spark와 같은 다른 애플리케이션을 클러스터에서 실행하지 않는 것이 좋습니다.

Amazon EMR 클러스터의 프라이머리 노드에 연결하고 쿼리를 실행합니다.

테스트 데이터 프로비저닝 및 권한 구성

Glue Data Catalog 및 해당 Hive AWS 메타스토어를 사용하여 Trino로 Amazon EMR을 테스트할 수 있습니다. 이러한 사전 조건 단계에서는 테스트 데이터를 설정하지 않은 경우 설정 방법을 설명합니다.

1. 아직 생성하지 않은 경우 통신 암호화에 사용할 SSH 키를 생성합니다.
2. 여러 파일 시스템 중에서 선택하여 데이터 및 로그 파일을 저장할 수 있습니다. 시작하려면 Amazon S3 버킷을 생성합니다. 버킷에 고유한 이름을 지정합니다. 생성할 때 생성한 암호화 키를 지정합니다.

   > **Note**
   >
   > 동일한 리전을 선택하여 스토리지 버킷과 Amazon EMR 클러스터를 모두 생성합니다.

3. 생성한 버킷을 선택합니다. 폴더 생성을 선택하고 폴더에 기억하기 쉬운 이름을 지정합니다. 폴더를 생성할 때 보안 구성을 선택합니다. 상위에 대한 보안 설정을 선택하거나 보안 설정을 더 전문화할 수 있습니다.
4. 폴더에 테스트 데이터를 추가합니다. 이 자습서에서는 쉼표로 구분된 레코드의 .csv를 사용하면이 사용 사례를 완료하는 데 효과적입니다.

5. Amazon S3 버킷에 데이터를 추가한 후 데이터 쿼리를 위한 추상화 계층을 제공하도록 AWS Glue에서 테이블을 구성합니다.

쿼리 연결 및 실행

다음은 Trino를 실행하는 클러스터에 연결하고 쿼리를 실행하는 방법을 설명합니다. 이렇게 하기 전에 메타스토어 테이블이 표시되도록 이전 절차에서 설명한 Hive 메타스토어 커넥터를 설정해야 합니다.

1. EC2 Instance Connect를 사용하여 클러스터에 연결하는 것이 좋습니다. 보안 연결을 제공하기 때문입니다. 클러스터 요약에서 SSH를 사용하여 프라이머리 노드에 연결을 선택합니다. 연결을 사용하려면 서브넷의 클라이언트에 대한 포트 22를 통한 연결을 허용하는 인바운드 규칙이 보안 그룹에 있어야 합니다. 또한 연결할 때 사용자 hadoop을 사용해야 합니다.
2. 를 실행하여 Trino CLI를 시작합니다`trino-cli`. 이렇게 하면 Trino를 사용하여 명령을 실행하고 데이터를 쿼리할 수 있습니다.
3. `show catalogs;`을(를) 실행합니다. Hive 카탈로그가 나열되어 있는지 확인합니다. 이는 데이터 스토어 또는 시스템 설정이 포함된 사용 가능한 카탈로그 목록을 제공합니다.
4. 사용 가능한 스키마를 보려면를 실행합니다`show schemas in hive;`. 여기에서 스키마의 이름을 실행`use schema-name;`하고 포함할 수 있습니다. 그런 다음를 실행`show tables;`하여 테이블을 나열할 수 있습니다.
5. 스키마의 테이블 이름을 `SELECT * FROM table-name`사용하여와 같은 명령을 실행하여 테이블을 쿼리합니다. 특정 스키마에 연결하기 위해 USE 문을 이미 실행한 경우 *schema.table*과 같은 두 부분으로 구성된 표기법을 사용할 필요가 없습니다.

Amazon EMR에서 Trino 구성

주제

- Trino용 커넥터 구성
- 모니터링

Trino용 커넥터 구성

Hive AWS 메타스토어로 Glue에 연결

Trino를 사용하여 쿼리를 실행할 때 AWS Glue 데이터 카탈로그를 Hive 메타스토어로 구성할 수 있다는 점을 이해하는 것이 중요하고 유용합니다. Hive 메타스토어로 클러스터를 설정하는 단계를 포함한 추가 정보는 Glue 데이터 카탈로그를 Hive의 AWS 메타스토어로 사용을 참조하세요.

EMR on EKS를 AWS Glue와 통합하는 방법에 대한 자세한 내용은 다음 모범 사례인 EMR Containers를 AWS Glue와 통합을 참조하세요.

Amazon EMR에서 Trino를 사용할 때 Iceberg 테이블에 연결

Iceberg는 분석 테이블의 오픈 테이블 형식입니다. Spark 및 Trino와 같은 엔진이 SQL 쿼리를 사용하여 동일한 테이블에서 빅 데이터를 쿼리하도록 생성되었습니다. 여기에는 데이터 읽기 및 쓰기 격리와 같은 기능이 포함되어 있으므로 리더는 예를 들어 부분적으로 업데이트된 데이터를 쿼리하지 않아도 됩니다. 스냅샷과 같은 상태 기능도 지원합니다. 메타데이터 및 매니페스트 파일을 사용하여 추상화 계층을 제공합니다. 테이블 스키마를 설명하고 형식 지정 또는 구성 방법에 대한 많은 세부 정보를 알 필요 없이 데이터를 쉽게 쿼리할 수 있습니다. 연결된 경우 테이블 업데이트 데이터에서 데이터를 읽거나 기본 파일에 새 데이터를 쓸 수 있습니다.

Amazon EMR 및 AWS Glue를 사용하여 Iceberg 테이블을 구성하는 방법을 보여주는 워크숍이 있습니다. 자세한 내용은 분석 워크숍 - Data Lake에서 Apache Iceberg 테이블 설정 및 사용을 참조하세요.

클라이언트와 연결

사용 가능한 JDBC 드라이버를 사용하여 Trino에 연결할 수 있습니다. 자세한 내용은 Trino 설명서의 JDBC 드라이버를 참조하세요.

모니터링

를 통해 Amazon EMR 클러스터를 모니터링할 수 있습니다 AWS Management Console. 자세한 내용은 Amazon EMR 클러스터가 작업을 수행할 때 보기 및 모니터링을 참조하세요. Amazon EMR은 모니터링 지표도 전송합니다 Amazon CloudWatch. Amazon EMR 클러스터 모니터링에 대한 자세한 내용은 Amazon CloudWatch Amazon EMR의 이벤트 및 지표를 참조하세요.

Amazon EMR의 Trino 모범 사례

Trino의 아키텍처는 각 구성 요소가 쿼리 실행에 특별한 역할을 하는 코디네이터-작업자 모델에 따라 여러 데이터 소스의 대규모 데이터 세트에 대한 빠른 분산 SQL 쿼리를 위해 설계되었습니다. 최상의

성능을 위해 Trino를 실행하는 Amazon EMR 클러스터를 구성하기 위해 집중할 수 있는 몇 가지 영역 또는 범주가 있습니다. 여기에는 다음이 포함됩니다.

- 메모리 최적화를 위한 클러스터 구성 설정 조정.
- 데이터 파티셔닝 및 데이터 배포를 위한 설정 최적화.
- 동적 필터링을 사용하여 쿼리 결과 수를 줄입니다.

이러한 설정 중 일부는 Amazon EMR에서 Trino를 사용할 때 자동으로 조정됩니다. 콘솔 또는 CLI 명령을 통해 수동으로 설정할 수 있습니다. 이 섹션의 주제는 데이터와 클러스터를 최적으로 구성하는 데 도움이 됩니다.

주제

성능 개선을 위한 주요 중점 영역

Trino는 쿼리 병렬 처리 및 메모리 최적화를 극대화합니다. 이 아키텍처는 효율적으로 규모를 조정하면서 다양한 여러 데이터 소스를 쿼리할 수 있도록 하여 유연성을 제공합니다. Trino에서 성능 개선의 주요 영역에는 아래 나열된 영역이 포함됩니다.

메모리 최적화

Trino의 메모리 관리는 특히 크고 복잡한 쿼리를 실행할 때 고성능과 안정성을 달성하는 데 매우 중요합니다. Trino는 분산 메모리 모델을 사용합니다. 이 모델에서는 작업, 집계, 조인 및 기타 작업을 처리하기 위해 작업자 노드에 메모리가 할당됩니다. 다음 목록은 이러한 설정 모음을 소개합니다.

- query.max-memory – 전체 클러스터에서 단일 쿼리에 사용할 수 있는 최대 메모리를 설정합니다. 이는 하드 제한입니다. 쿼리가이 메모리를 초과하면 실패합니다.
- query.max-memory-per-node - 쿼리가 각 작업자 노드에서 사용할 수 있는 최대 메모리를 정의합니다. 이를 설정하면 단일 쿼리가 작업자의 리소스를 독점하지 않습니다.
- JVM 힙 크기 - JVM 수준에서 구성되며 각 노드에서 Trino 서버 프로세스의 최대 힙 크기를 설정합니다. 이 값은 일반적으로 JVM 수준에서 시스템의 메모리 부족을 방지하기 위해 Trino의 메모리 관련 구성(max-memory-per-node와 heap-headroom-per-node의 합계)보다 커야 합니다.

- memory.heap-headroom-per-node - 쿼리가 아닌 작업에 대해 JVM 힙 크기를 벗어나도록 메모리의 버퍼 양을 지정합니다. 이는 내부 작업 및 가비지 수집을 위한 충분한 오버헤드를 보장하는 데 매우 중요합니다.

동적 필터링

Trino의 동적 필터링은 특히 조인 중에 처리되는 데이터의 양을 줄여 쿼리 성능을 개선하는 최적화 기법입니다. 필터 조건을 동적으로 적용하여 다른 쪽에서 볼 수 있는 데이터를 기반으로 조인의 한쪽에서 스캔하는 데이터를 제한합니다. 이는 조인의 한쪽이 매우 선택적인(즉, 작은 데이터 하위 집합을 포함하는) 쿼리에 특히 유용합니다. Amazon EMR에서 기본적으로 활성화됩니다. 다음은 쿼리의 예입니다.

```
SELECT orders.order_id, orders.total_amount
FROM orders
JOIN customers ON orders.customer_id = customers.customer_id
WHERE customers.country = 'France';
```

동적 필터링이 없으면 Trino는 조인에서 전체 주문 테이블을 스캔하지만, 일부 고객(프랑스)만 관련이 있습니다. 이 접근 방식은 주문 테이블의 모든 행을 읽으므로 I/O 및 처리 비용이 높습니다. 동적 필터링을 사용하면 Trino는 처음에 작은 고객 테이블을 스캔하고 프랑스의 고객에 대해서만 customer_id 값을 검색한 다음이 하위 집합을 주문에 필터로 적용합니다. 즉, customer_id가 필터링된 하위 집합과 일치하는 주문의 관련 행만 스캔되므로 처리된 레코드가 크게 줄어듭니다.

디스크에 유출

Trino에서 디스크 유출을 사용하면 중간 쿼리 결과를 디스크로 오프로드할 수 있으므로 `query_max_memory` 또는에서 설정한 메모리 제한을 초과하는 경우에도 메모리 집약적인 쿼리를 완료할 수 있습니다`query_max_memory_per_node`. 기본적으로 Trino는 이러한 제한을 적용하여 공정한 메모리 할당을 보장하고 클러스터 교착 상태를 방지합니다. 그러나 대규모 쿼리가 이러한 제한을 초과하면 종료될 위험이 있습니다. 디스크 유출은를 사용하여이 문제를 해결`revocable memory`하므로 다른 곳에서 리소스가 필요한 경우 취소할 수 있는 추가 메모리를 쿼리에서 빌릴 수 있습니다. 메모리가 취소되면 중간 데이터가 디스크로 유출되어 메모리 제한을 초과하지 않고 쿼리를 계속 처리할 수 있습니다. 강제로 디스크로 유출되는 쿼리는 실행 시간이 길어질 수 있으므로 기본적으로 비활성화되어 있습니다. Amazon EMR에서 유출을 활성화하려면 다음 구성을 사용합니다.

- `spill-enabled=true` - 메모리 사용량이 사용 가능한 임계값을 초과할 때 디스크 유출을 활성화합니다.
- `spill-paths` - 유출된 데이터가 저장되는 디렉터리를 정의합니다`spill-paths=/mnt/spill`.

테이블 통계 수집 및 활용

테이블 통계를 수집하면 Trino의 비용 기반 최적화 프로그램이 조인 주문, 필터 푸시다운 및 파티션 정리에 대해 정보에 입각한 결정을 내릴 수 있으므로 성능이 향상됩니다.

ANALYZE 명령을 사용하여 Hive 또는 Iceberg 테이블에 대한 통계를 수집할 수 있습니다.

```
ANALYZE sales;
```

넓은 테이블에 대한 통계를 수집하면 리소스에 대한 세금이 부과될 수 있습니다. 조인, 필터 또는 그룹화 작업에 사용되는 열의 하위 집합을 지정하는 것이 좋습니다.

이 명령은 또 다른 유용한 명령입니다. 테이블에 대한 현재 통계를 표시하여 통계가 최신 상태인지 확인합니다.

```
show stats for table_name;
```

Trino 워크로드 확장 시 일반적인 문제

Trino와 함께 Amazon S3를 사용할 때 얻을 수 있는 주요 이점은 대용량 데이터 볼륨에 맞게 확장할 수 있는 S3의 기능과 S3의 비용 효율성입니다. 그러나 대용량 데이터 볼륨을 쿼리할 때 관련 성능 문제 모음이 가끔 발생할 수 있습니다. 이는 데이터가 저장되는 방식이나 우수한 성능을 제한하는 구성 설정 또는 기타 이유로 인해 발생할 수 있습니다. 이러한 문제가 발생하면 이를 방지하거나 완화하기 위해 취할 수 있는 효과적인 단계가 있습니다.

이 섹션은 대용량 데이터 볼륨에서 쿼리 성능을 높이기 위해 구현할 수 있는 일반적인 최적화 목록으로 시작합니다. 그런 다음 일반적인 문제가 자세히 설명되고 각 문제에 대한 완화 조치가 제공됩니다.

이 주제는 대규모 성능 가속화: Amazon S3를 사용한 Trino 모범 사례 컨퍼런스 프레젠테이션에서 발췌했습니다.

대규모 데이터 세트에 대한 데이터 레이아웃 최적화

대규모 데이터 세트를 쿼리할 때는 성능 병목 현상이 자주 발생하지 않습니다. 하지만 Trino를 사용하여 Amazon S3의 데이터를 쿼리할 때 헤드 스타트를 제공하기 위해 구현할 수 있는 모범 사례가 있습니다. 여기에는 다음이 포함됩니다.

- 파티셔닝 - 파티셔닝은 계층 구조에서 데이터를 구성하고 관련 속성을 기반으로 관련 데이터를 함께 저장하는 것을 의미합니다. 파티셔닝을 사용하면 쿼리가 관련 없는 데이터를 많이 스캔할 필요가 없으므로 쿼리 성능이 향상됩니다. 소스 데이터를 접두사, 특히 날짜 범위 리전 또는 기타 속성으로 정

렬하는 등 다양한 파티셔닝 전략을 사용할 수 있습니다. 성능을 높이기 위해 Amazon S3에서 데이터를 파티셔닝하는 방법에 대한 자세한 내용은 블로그 게시물 AWS Glue 데이터 카탈로그에서 지원하는 Amazon S3 테이블의 파티션 관리를 시작하기 또는 상위 10개 성능 튜닝 팁 게시물을 참조하세요 Amazon Athena.

- 버킷팅 - 버킷팅은 관련 데이터를 공통 파일로 그룹화합니다. 예를 들어 상태와 같은 지리적 리전에 따라 데이터를 쿼리하는 경우 동일한 파일 또는 파일 그룹에서 특정 상태에 대한 모든 데이터를 그룹화하여 쿼리 성능을 높일 수 있습니다. 이를 가장 잘 수행하려면 예를 들어 주 또는 도와 같이 카디널리티가 높은 데이터 속성을 기반으로 버킷팅합니다. 또한 쿼리 패턴을 고려할 수 있습니다. 쿼리가 일반적으로 해당 상태의 데이터를 함께 읽는 경우 캘리포니아와 오레곤에 대한 데이터를 그룹화하는 것을 예로 들 수 있습니다.
- S3 접두사 관리 - Amazon S3 접두사를 사용하여 파티셔닝 전략을 구현할 수 있습니다. 예를 들어 특정 날짜와 같은 Amazon S3 버킷에 대해 단일 접두사만 사용하는 경우 요청 수가 많아지고 HTTP 503 오류가 발생할 수 있습니다. 접두사를 사용하여 조건을 추가하고 소스 데이터를 보다 효과적으로 구성하는 것이 좋습니다. 자세한 내용은 Amazon S3 설명서의 접두사를 사용하여 객체 구성을 참조하세요. 다음 간단한 예제는 요청 처리량을 높이는 접두사를 보여줍니다`s3://bucket/country=US/dt=2024-06-13`. 이 샘플에서는 국가와 날짜가 모두 접두사에 포함되므로 접두사에 날짜만 포함된 경우보다 읽기 수가 적습니다.

  HTTP 503 오류 완화는이 주제에서 이어지는 HTTP 속도 저하 섹션에서 자세히 설명합니다.
- 데이터 크기 최적화 - OPTIMIZE 명령을 실행하여 성능이 더 좋은 쿼리에 도움이 되는 구성을 설정할 수 있습니다. Hive 외부 테이블에 대해 실행하려면 다음 단계를 따릅니다.
  - 를 파라미터와 `OPTIMIZE` 함께 사용합니다`hive.non-managed-table-writes-enabled=true`. 이 속성에 대한 자세한 내용은 Hive 일반 구성 속성을 참조하세요.
  - 다음 세션 파라미터를 설정합니다. `SET SESSION catalog.non_transactional_optimize_enabled=true`
  - `OPTIMIZE` 명령을 실행합니다`ALTER TABLE catalog.schema.table EXECUTE optimize(file_size_threshold => '128MB')`. 이 경우 `file_size_threshold`는 기본적으로 100MB입니다. 샘플에 표시된 대로이 임계값을 높이면 128MB 미만의 파일이 병합됩니다.
- 재시도 구성 -를 설정하여 HTTP 503 오류 가능성을 완화할 수 있는 재시도 제한을 늘릴 수 있습니다`s3.max-error-retries`. 이는 TrinoFileSystem API 및 Trino 449 버전 이상을 사용할 때 적용됩니다. 반면 Trino와 함께 Amazon EMR을 사용하는 경우 EMRFS를 사용하여 Amazon S3에 액세스합니다. EMRFS를 사용하면 `fs.s3.maxRetries` 파라미터를 변경하여 사용 중지 수를 늘릴 수 있습니다.
- Amazon S3 스토리지 클래스 선택 - 수명 주기의 여러 지점에서 데이터에 적합한 스토리지 클래스를 선택하면 특정 데이터 수집에 대한 요구 사항에 따라 성능과 비용 모두에 도움이 될 수 있습니

다. 자세한 내용은 Amazon S3 설명서의 Amazon S3 스토리지 클래스 이해 및 관리를 참조하세요. Amazon S3

- Iceberg로 마이그레이션 - 성능 문제, 특히 작은 파일에서 쿼리를 실행하는 것과 관련된 문제를 완화하는 또 다른 솔루션은 Iceberg 테이블로 마이그레이션하는 것입니다. Iceberg에는 작은 파일을 잘 처리하는 기능이 있습니다.
- 자동 데이터 압축 사용 - Iceberg 테이블을 사용하는 경우 AWS Glue 데이터 카탈로그를 사용한 자동 데이터 압축은 데이터 크기를 최적화하고 쿼리 성능을 개선할 수 있습니다.

대규모 데이터 세트를 쿼리할 때 발생하는 일반적인 문제

이 섹션에서는 Amazon S3에 대용량 데이터 세트를 누적하고 Trino로 쿼리할 때 발생할 수 있는 일반적인 문제 모음을 나열합니다. 각 섹션에서는 문제를 해결하거나 쿼리에 미치는 영향을 줄이는 방법을 보여줍니다. 다음 섹션에 설명된 각 문제는 Hive 커넥터를 사용하여 재현 및 테스트되었습니다.

대규모 데이터 스캔

쿼리가 대용량 데이터 세트를 스캔해야 하는 경우 쿼리 성능 저하 및 스토리지 비용 상승과 같은 문제가 발생할 수 있습니다. 대규모 데이터 볼륨은 적절한 기간 내에 레거시 데이터를 이동하지 않는 빠른 데이터 증가 또는 계획으로 인해 발생할 수 있습니다. 이로 인해 쿼리 속도가 느려질 수 있습니다.

대용량 데이터 세트 스캔으로 인한 성능 저하를 완화하려면 파티셔닝 및 버킷팅을 활용하는 것이 좋습니다.

- 파티셔닝은 속성을 기반으로 관련 데이터를 그룹화합니다. 파티셔닝을 효과적으로 사용하면 쿼리 성능이 크게 향상될 수 있습니다.
- 버킷팅은 특정 관련 데이터 열에 따라 파일 또는 버킷의 데이터를 그룹화하는 것을 말합니다. 버킷팅은 일반적으로 관련 소스 데이터 파일을 물리적으로 함께 유지하는 것을 의미합니다.

대규모 데이터 스캔에서 완화가 어떻게 작동하는지 설명하기 위해 캘리포니아 또는 알래스카에 할당할 수 있는 상태 속성이 있는 레코드가 있고이 상태 속성이 쿼리 조건 중 하나인 데이터를 저장하고 쿼리한다고 가정합니다. S3 접두사를 사용하여 각 상태에 대한 데이터를 별도의 S3 버킷에 저장하거나 상태에 따라 데이터를 파티셔닝하여 쿼리 성능을 개선할 수 있습니다. 이 파티셔닝 및 버킷팅은 날짜 속성과 같은 추가 열을 기반으로 하는 경우에도 성능을 개선할 수 있습니다.

> **Note**
> 열의 카디널리티가 높고 이를 사용하여 데이터를 그룹화하려는 경우이 경우 버킷팅을 사용하는 것이 좋습니다. 반면, 일반적으로 파티션 키는 카디널리티가 더 낮아야 합니다.

다양한 S3 스토리지 유형 사용

일반적으로 워크로드의 성능, 데이터 액세스, 복원력 및 비용 요구 사항을 기반으로 스토리지 유형을 선택합니다. 비용과 성능 간에 절충이 있을 수 있습니다. 데이터 액세스 패턴과 일치하는 적절한 Amazon S3 스토리지 클래스를 선택하는 것이 중요합니다. 두 가지 주요 액세스 패턴이 있습니다.

- 알려지거나 예측 가능한 방식으로 액세스되는 데이터입니다. 일반적으로 자주 액세스하지 않는 데이터가 있는 경우 비용을 절감하는 데 도움이 되므로 S3 Standard IA를 선택하는 것이 좋습니다. 데이터에 자주 액세스한 경우 S3 Standard는 Amazon EMR 및 Trino를 사용한 액세스에 가장 적합합니다.
- 알 수 없거나 예측할 수 없는 방식으로 액세스되는 데이터입니다. 이렇게 하면 다른 Amazon S3 스토리지 클래스를 사용하기 위해를 호출할 수 있습니다. S3 스토리지 클래스 간에 절충이 있습니다. 여기에는 지연 시간, 스토리지 비용 및 가용성이 포함됩니다. 워크로드 및 액세스 패턴에 따라 적절한 S3 스토리지 유형을 선택할 수 있습니다. 각 클래스의 이점에 대한 설명은 [Amazon S3 스토리지 클래스를 참조하세요](https://aws.amazon.com/s3/storage-classes/).

압축 사용

Iceberg 테이블을 사용하면 Iceberg 자동 압축을 사용하여 파일 크기를 최적화하여 쿼리 효율성을 높일 수도 있습니다. 자세한 내용은 이제 AWS Apache Iceberg 테이블의 자동 압축을 지원하는 Glue 데이터 카탈로그를 참조하세요.

HTTP 속도 저하 오류

이는 요청 속도가 Amazon S3 접두사에 대해 미리 구성된 임계값을 초과할 때 발생합니다. 이 상태에 도달할 때 가장 일반적으로 발생하는 HTTP 오류는 오류 503: 요청 속도를 줄이세요. 이 문제의 소스는 데이터를 읽기 위해 생성해야 하는 분할 수가 많기 때문에 작은 파일이 많을 때 루트가 될 수 있습니다. 문제를 완화하는 몇 가지 방법이 있습니다.

- Trino에서 Amazon S3 요청에 대한 재시도 제한을 늘립니다. 이는 Trino 449`fs.s3.maxretries`에서를 사용하는 EMRFS에 대해 설정됩니다.
- 파일 크기를 최적화하면 요청 속도가 낮아질 수도 있습니다.

Trino가 쿼리할 데이터 세트의 분할 수를 결정하는 방법에 대한 자세한 내용은 Hive 커넥터 설명서의 성능 튜닝 구성 속성을 참조하세요.

작은 파일을 쿼리하기 어려움

많은 작은 파일을 쿼리하면 GET 및 LIST 요청 수가 많아서 I/O 오버헤드가 높아지고 쿼리 성능에 부정적인 영향을 미칠 수 있습니다. 파일 크기를 최적화하면 쿼리 성능이 향상될 수 있습니다. 이를 수행하는 몇 가지 방법이 있습니다.

- 데이터를 더 적은 수의 더 큰 파일로 통합합니다. (일반적으로 파일 크기는 약 128MB로 유지하는 것이 좋습니다.) ETL 파이프라인에서와 같이 데이터를 수집할 때 도구를 사용하여이 작업을 수행하거나 데이터를 수동으로 통합할 수 있습니다. 이러한 솔루션을 사용할 수 없는 경우 나머지 옵션이 더 적합할 수 있습니다.
- `OPTIMIZE` 명령을 실행합니다.
- `SESSION` 파라미터를 설정합니다.

Iceberg에는 작은 파일을 자동 압축인 더 큰 파일로 병합할 수 있는 기능이 있습니다. AWS Glue 데이터 카탈로그로 관리되는 파일에서 작동합니다. 자세한 내용은 이제 AWS Glue Data Catalog에서 Apache Iceberg 테이블의 자동 압축 지원 섹션을 참조하세요.

필요하지 않은 데이터가 포함된 쿼리

데이터가 증가하는 것은 일반적이므로 데이터 액세스 패턴을 추적하고 데이터가 노후화되거나 관련이 없게 되면 데이터를 적절하게 이동해야 합니다. 이는 데이터가 증가함에 따라 쿼리 실행 시 스캔할 데이터의 양이 많기 때문에 시간이 지남에 따라 쿼리 성능이 저하될 수 있기 때문입니다. Amazon S3 및 기타 서비스는 데이터 수명 주기 마이그레이션에 대한 지침을 제공합니다.이 지침은 데이터를 다른 스토리지 위치로 이동하는 전략을 보여줍니다. 이렇게 하면 스토리지 비용 이점도 있습니다.

데이터 마이그레이션 외에도 실행 중인 쿼리와 관련이 없는 소스 데이터 제거와 같은 다른 전략을 사용할 수 있습니다. 이는 소스 데이터 스키마를 변경하는 것을 의미할 수 있으므로 일부 작업이 필요할 수 있습니다. 하지만 긍정적인 결과는 데이터 볼륨을 줄이고 쿼리 속도를 높이는 것입니다. 자세한 내용은 객체의 수명 주기 관리를 참조하세요.

Trino 고려 사항

Amazon EMR에서 Trino를 실행할 때 다음 사항을 고려하세요.

구성할 수 없는 Trino 배포 속성

다음 표에는 Trino `properties` 파일의 다양한 구성 옵션이 나와 있습니다.

파일	구성 가능
`log.properties`	Trino: Amazon EMR 버전 6.1.0 이상에서 구성 가능합니다. `prestosql-log` 또는 `trino-log` 구성 분류를 사용합니다.
`config.properties`	Trino: Amazon EMR 버전 6.1.0 이상에서 구성 가능합니다. `prestosql-config` 또는 `trino-config` 구성 분류를 사용합니다.
`hive.properties`	Trino: Amazon EMR 버전 6.1.0 이상에서 구성 가능합니다. `prestosql-connector-hive` 또는 `trino-connector-hive` 구성 분류를 사용합니다.
`node.properties`	Trino: Amazon EMR 버전 6.1.0 이상에서 구성 가능합니다. `prestosql-node` 또는 `trino-node` 구성 분류를 사용합니다.
`jvm.config`	구성할 수 없습니다.

추가 고려 사항

- EMR 버전 6.1.0 이상의 Trino의 경우 Amazon EMR은 클러스터 노드 간의 안전한 내부 통신을 위해 공유 보안 암호 키를 자동으로 구성합니다. 이 보안 기능을 활성화하기 위해 추가 구성을 수행할 필요가 없으며 자체 보안 암호 키를 사용하여 구성을 재정의할 수 있습니다. Trino 내부 인증에 대한 자세한 내용은 Trino 353 설명서: Secure internal communication을 참조하세요.

Trino 릴리스 기록

릴리스 정보 섹션에서는 Amazon EMR 기반 Trino의 특정 버전에 대한 변경 및 업데이트에 대해 자세히 설명합니다.

버전별 Trino 릴리스 정보

Amazon EMR 7.6.0 - Trino 릴리스 정보

Amazon EMR 7.6.0 - Trino 새 기능

- 이제 Trino에는 장기 실행 쿼리를 지원하기 위해 내결함성 실행 메커니즘이 포함됩니다. 내결함성 실행은 실패한 쿼리 또는 구성 요소 작업을 재시도하여 쿼리 실패를 완화합니다.

Amazon EMR 7.6.0 - Trino 변경 사항

Amazon EMR 7.6.0 - Trino 변경 사항

유형	설명
업그레이드	457로 Trino 업그레이드

Amazon EMR 7.3.0 - Trino 릴리스 정보

Amazon EMR 7.3.0 - Trino 변경 사항

- 이 릴리스에서는 Trino를 버전 436에서 442로 업그레이드합니다.
- 이 릴리스에서는 Hudi 쿼리를 새 Hudi 보정기로 리디렉션합니다. 이전 Hive 커넥터는 더 이상 Hudi 테이블을 읽을 수 없습니다. Note
- 이 릴리스에서는 Rubix 모듈이 이제 오픈 소스에서 더 이상 사용되지 않으므로 Amazon EMR에서 Rubix 모듈을 제거합니다.
- 이 릴리스에서는 `hive.security` 속성에서 레거시 모드를 제거합니다. 이제 기본값은 `allow-all`입니다.

Amazon EMR 6.9.0 - Trino 릴리스 정보

Amazon EMR 6.9.0 - 새로운 기능 Trino

- 이제 Trino에는 장기 실행 쿼리를 지원하기 위해 내결함성 실행 메커니즘이 포함됩니다. 내결함성 실행은 실패한 쿼리 또는 구성 요소 작업을 재시도하여 쿼리 실패를 완화합니다.

Amazon EMR 6.9.0 - Trino 변경

Amazon EMR 6.9.0 - Trino 변경

유형	설명
업그레이드	Trino를 398로 업그레이드
업그레이드	Hadoop 3.3.3에 대한 지원
Feature	Tardigrade 지원: HDFS 및 Amazon S3에서 교환 스풀링에 대한 지원을 추가합니다.
버그 수정	Trino Iceberg를 사용하고 Glue 카탈로그를 활성화한 경우 `iceberg.properties.` 에서 메타스토어 URI를 추가하지 않습니다.

Amazon EMR 6.9.0 - 알려진 문제 Trino

- Amazon EMR 릴리스 6.9.0의 경우 Trino는 Apache Ranger가 활성화된 클러스터에서 작동하지 않습니다. Ranger와 함께 Trino를 사용해야 하는 경우 지원에 문의하세요.

Apache Zeppelin

대화식 데이터 탐색을 가능하게 하는 노트북으로 Apache Zeppelin을 사용하십시오. Zeppelin에 대한 자세한 내용은 https://zeppelin.apache.org/를 참조하세요. Zeppelin은 Amazon EMR 릴리스 5.0.0 이상에 포함되어 있습니다. 이전 릴리스에서는 Zeppelin이 샌드박스 애플리케이션으로 포함되었습니다. 자세한 내용은 Amazon EMR 4.x 릴리스 버전 단원을 참조하십시오.

Zeppelin 웹 인터페이스에 액세스하려면 마스터 노드와 프록시 연결에 대한 SSH 터널을 설정하십시오. 자세한 내용은 EMR 클러스터에 호스팅된 웹 인터페이스 보기를 참조하세요.

다음 테이블에는 Amazon EMR이 Zeppelin을 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 Zeppelin의 버전이 나열되어 있습니다.

이 릴리스에서 Zeppelin과 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 Zeppelin 버전 정보

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-7.9.0	Zeppelin 0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

다음 테이블에는 Amazon EMR이 Zeppelin을 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 Zeppelin의 버전이 나열되어 있습니다.

이 릴리스에서 Zeppelin과 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 Zeppelin 버전 정보

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.15.0	Zeppelin 0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

다음 테이블에는 Amazon EMR이 Zeppelin을 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 Zeppelin의 버전이 나열되어 있습니다.

이 릴리스에서 Zeppelin과 함께 설치된 구성 요소의 버전은 릴리스 5.36.2 구성 요소 버전을 참조하세요.

emr-5.36.2용 Zeppelin 버전 정보

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.36.2	Zeppelin 0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
		httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

주제

- Amazon EMR에서 Zeppelin 사용 시 고려 사항
- Zeppelin 릴리스 기록

Amazon EMR에서 Zeppelin 사용 시 고려 사항

- 마스터 노드의 다른 웹 서버에 연결하려면 동일한 SSH 터널링 메서드를 사용하여 Zeppelin에 연결합니다. Zeppelin 서버는 포트 8890에 있습니다.
- Amazon EMR 릴리스 버전 5.0.0 이상에서 Zeppelin은 Shiro 인증을 지원합니다.
- Amazon EMR 릴리스 버전 5.8.0 이상의 Zeppelin은 Glue 데이터 카탈로그를 Spark SQL용 AWS 메타스토어로 사용할 수 있도록 지원합니다. 자세한 내용은 Glue 데이터 카탈로그를 Spark SQL AWS용 메타스토어로 사용을 참조하세요.
- Zeppelin은 클러스터의 `spark-defaults.conf` 구성 파일에 정의된 일부 설정을 사용하지 않습니다. 이는 `spark.dynamicAllocation.enabled`를 `true`로 설정한 경우 YARN에게 실행기를 동적으로 할당하도록 지시한 경우에도 마찬가지입니다. Zeppelin Interpreter(Zeppelin 인터프리터) 탭을 사용하여 실행기 설정(예: 메모리 및 코어)을 설정한 후 인터프리터를 다시 시작해야 이 설정을 사용할 수 있습니다.
- Amazon EMR 릴리스 6.10.0 이상에서는 Apache Flink와의 Apache Zeppelin 통합을 지원합니다. 자세한 정보는 Amazon EMR의 Zeppelin에서 Flink 작업을 참조하세요.
- Amazon EMR에서 Zeppelin은 SparkR 인터프리터를 지원하지 않습니다.

Zeppelin 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 Zeppelin의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

Zeppelin 버전 정보

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-7.9.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.8.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server,

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
		spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.7.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.6.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-7.5.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.4.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-7.3.0	0.11.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.2.0	0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.36.2	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-7.1.0	0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-7.0.0	0.10.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.15.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.14.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.13.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.12.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.11.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.11.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.10.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.10.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.9.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.9.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.8.1	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.8.0	0.10.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.7.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.36.1	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.36.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.6.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, hudi, hudi-spark, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.35.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.5.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.4.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.3.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.3.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.2.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.2.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.1.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.1.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-6.0.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-6.0.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.34.0	0.10.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.33.1	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.33.0	0.9.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.32.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.32.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.31.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.31.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.30.2	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.30.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.30.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.29.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.28.1	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.28.0	0.8.2	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.27.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.27.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.26.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.25.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.24.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.24.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.23.1	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.23.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.22.0	0.8.1	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, livy-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.21.2	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.21.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.21.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.20.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.20.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.19.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.19.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.18.1	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.18.0	0.8.0	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.17.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.17.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.17.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.16.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.16.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.15.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.15.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.14.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.14.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.14.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.13.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.13.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, r, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.12.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.12.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.12.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.4	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.11.3	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.2	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.11.1	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.11.0	0.7.3	aws-sagemaker-spark-sdk, emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.10.1	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.10.0	0.7.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.9.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.9.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.8.3	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.2	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.8.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.8.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.7.1	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.7.0	0.7.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.6.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.6.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.5.4	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.3	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.5.2	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.5.1	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.5.0	0.7.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.4.1	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.4.0	0.7.0	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.3.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.3.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.3.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.2.3	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.2	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.2.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.2.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.1.1	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.1.0	0.6.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.0.3	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.2	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Amazon EMR 릴리스 레이블	Zeppelin 버전	Zeppelin과 함께 설치된 구성 요소
emr-5.0.1	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server
emr-5.0.0	0.6.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, spark-client, spark-history-server, spark-on-yarn, spark-yarn-slave, zeppelin-server

Apache ZooKeeper

Apache ZooKeeper는 구성 정보 유지, 명명, 분산 동기화 제공 및 그룹 서비스 제공을 위한 중앙화된 서비스입니다. ZooKeeper에 대한 자세한 내용은 http://zookeeper.apache.org/를 참조하십시오.

다음 테이블에는 Amazon EMR이 ZooKeeper를 통해 설치하는 구성 요소와 함께 Amazon EMR 7.x 시리즈의 최신 릴리스에 포함된 ZooKeeper의 버전이 나열되어 있습니다.

이 릴리스에서 ZooKeeper와 함께 설치된 구성 요소의 버전은 릴리스 7.9.0 구성 요소 버전을 참조하세요.

emr-7.9.0용 ZooKeeper 버전 정보

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-7.9.0	ZooKeeper 3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

다음 테이블에는 Amazon EMR이 ZooKeeper를 통해 설치하는 구성 요소와 함께 Amazon EMR 6.x 시리즈의 최신 릴리스에 포함된 ZooKeeper의 버전이 나열되어 있습니다.

이 릴리스에서 ZooKeeper와 함께 설치된 구성 요소의 버전은 릴리스 6.15.0 구성 요소 버전을 참조하세요.

emr-6.15.0용 ZooKeeper 버전 정보

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.15.0	ZooKeeper 3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

다음 테이블에는 Amazon EMR이 ZooKeeper를 통해 설치하는 구성 요소와 함께 Amazon EMR 5.x 시리즈의 최신 릴리스에 포함된 ZooKeeper의 버전이 나열되어 있습니다.

이 릴리스에서 ZooKeeper와 함께 설치된 구성 요소의 버전은 [릴리스 5.36.2 구성 요소 버전]을 참조하세요.

emr-5.36.2용 ZooKeeper 버전 정보

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.36.2	ZooKeeper 3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

주제

- ZooKeeper 릴리스 기록
- 버전별 ZooKeeper 릴리스 정보

ZooKeeper 릴리스 기록

다음 테이블에는 애플리케이션과 함께 설치된 구성 요소 외에도 Amazon EMR의 최신 릴리스에 포함된 ZooKeeper의 버전이 나열되어 있습니다. 각 릴리스의 구성 요소 버전은 Amazon EMR 7.x 릴리스 버전, Amazon EMR 6.x 릴리스 버전 또는 Amazon EMR 5.x 릴리스 버전의 릴리스의 구성 요소 버전 섹션을 참조하세요.

ZooKeeper 버전 정보

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-7.9.0	3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.8.0	3.9.3	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-7.7.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.6.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.5.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-7.4.0	3.9.2	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.3.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.2.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.36.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.1.0	3.9.1	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-7.0.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.15.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.14.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.13.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.12.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.11.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.10.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.10.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.9.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.9.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.1	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.8.0	3.5.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.7.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.36.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.6.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.35.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.5.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.4.0	3.5.7	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.3.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.2.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.2.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.1.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-6.1.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-6.0.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.34.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.33.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.32.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.32.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.31.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.31.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.2	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.30.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.30.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.29.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.28.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.28.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.1	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.27.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.26.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.25.0	3.4.14	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.24.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.24.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.23.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.22.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.2	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.21.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.21.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.20.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.19.1	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.19.0	3.4.13	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.18.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.18.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.2	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.17.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.17.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.16.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.15.1	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.15.0	3.4.12	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.14.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.14.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.13.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.13.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.12.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.12.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.11.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.11.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.11.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.10.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.9.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.9.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.8.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.8.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.7.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.7.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.6.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.6.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, hadoop-yarn-timeline-server, zookeeper-client, zookeeper-server
emr-5.5.4	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.3	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.5.2	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.1	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.5.0	3.4.10	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.4.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.4.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.3.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.3.0	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.3	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.2.2	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.1	3.4.9	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.2.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.1.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.1.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.3	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

Amazon EMR 릴리스 레이블	ZooKeeper 버전	ZooKeeper와 함께 설치된 구성 요소
emr-5.0.2	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.1	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server
emr-5.0.0	3.4.8	emrfs, emr-goodies, hadoop-client, hadoop-hdfs-datanode, hadoop-hdfs-library, hadoop-hdfs-namenode, hadoop-httpfs-server, hadoop-kms-server, hadoop-yarn-nodemanager, hadoop-yarn-resourcemanager, zookeeper-client, zookeeper-server

버전별 ZooKeeper 릴리스 정보

전체 릴리스 정보는 다음 섹션을 참조하세요.

Amazon EMR 7.8.0 - ZooKeeper 릴리스 정보

Amazon EMR 7.8.0 - ZooKeeper 변경 사항

업그레이드 - Zookeeper가 버전 3.9.3으로 업그레이드되었습니다.

Amazon EMR 7.5.0 - ZooKeeper 릴리스 정보

Amazon EMR 7.5.0 - ZooKeeper 변경 사항

- EMR-7.5.0부터 Zookeeper는 런타임 시 기본적으로 Java 17로 설정됩니다. Java 런타임에 다른 버전을 사용하려면 `zookeeper-server` 파일의 JVM 설정을 재정의하고 `JAVA_HOME` 원하는 버전으로 설정합니다.

Amazon EMR 7.4.0 - ZooKeeper 릴리스 정보

유형	설명
업그레이드	Zookeeper 버전이 3.9.2로 업그레이드되었습니다.
개선	ZOOKEEPER-4778: 심각도가 높은 취약성을 제거하기 위해 jetty, netty 및 logback을 패치합니다.
개선	ZOOKEEPER-4799: CVE-2024-23944 수정.
버그 수정	ZOOKEEPER-4415: TLSv1.3 지원이 추가되었습니다.

Amazon EMR 7.4.0 - ZooKeeper 기능

- 전송 중 암호화 보안 구성이 제공된 클러스터는 기본적으로 Quorum 통신 및 Zookeeper Admin Server에 대해 TLS가 활성화됩니다. 또한 TLS 보안 포트는 Zookeeper와의 클라이언트 통신을 위해 2281에 노출됩니다.

커넥터 및 유틸리티

Amazon EMR은 다른 AWS 서비스에 데이터 소스로 액세스할 수 있는 여러 커넥터와 유틸리티를 제공합니다. 일반적으로 프로그램 내에서 이러한 서비스의 데이터에 액세스할 수 있습니다. 예를 들면 Hive 쿼리, Pig 스크립트 또는 MapReduce 애플리케이션에서 Kinesis 스트림을 지정하여 이러한 데이터에 대한 작업을 수행할 수 있습니다.

주제

Amazon EMR을 사용하여 DynamoDB에서 테이블 내보내기, 가져오기, 쿼리 및 조인

> **Note**
>
> Amazon EMR-DynamoDB 커넥터는 GitHub에서 오픈 소스로 제공됩니다. 자세한 내용은 https://github.com/awslabs/emr-dynamodb-connector를 참조하십시오.

DynamoDB는 완전관리형 NoSQL 데이터베이스 서비스로서 원활한 확장성과 함께 빠르고 예측 가능한 성능을 제공합니다. 개발자는 데이터베이스 테이블을 만들고 해당 요청 트래픽 또는 스토리지를 제한 없이 확장할 수 있습니다. DynamoDB는 테이블의 데이터와 트래픽을 충분한 수의 서버로 자동 분산하여 고객이 지정한 요청 용량과 저장된 데이터 규모를 처리하면서도 일관되고 빠른 성능을 발휘합니다. 여기에 Amazon EMR 및 Hive를 사용하면 DynamoDB에 저장된 데이터와 같은 대용량 데이터도 빠르고 효율적으로 처리할 수 있습니다. DynamoDB에 대한 자세한 내용은 Amazon DynamoDB 개발자 안내서를 참조하세요.

Apache Hive는 사용이 쉬우면서 SQL과 유사한 쿼리 언어(HiveQL)로 map reduce 클러스터에 대해 쿼리를 실행할 때 사용하는 소프트웨어 계층입니다. 이 소프트웨어는 하둡 아키텍처를 기반으로 실행됩니다. Hive 및 HiveQL에 대한 자세한 내용은 HiveQL language manual을 참조하세요. Hive 및 Amazon EMR에 대한 자세한 내용은 Apache Hive 섹션을 참조하세요.

DynamoDB 연결 기능이 포함된 사용자 지정 버전의 Hive와 함께 Amazon EMR 사용하여 DynamoDB에 저장된 데이터에 대해 다음 작업을 수행할 수 있습니다.

- DynamoDB 데이터를 Hadoop 분산 파일 시스템(HDFS)에 로드한 후 Amazon EMR 클러스터에 대한 입력으로 사용.
- 유사 SQL 문(HiveQL)을 사용하여 실시간 DynamoDB 데이터 쿼리.
- DynamoDB에 저장된 데이터 조인 및 내보내기 또는 조인된 데이터 쿼리.
- DynamoDB에 저장된 데이터를 Amazon S3로 내보내기.
- Amazon S3에서 DynamoDB로 데이터 가져오기.

> **Note**
>
> Amazon EMR-DynamoDB 커넥터는 Kerberos 인증을 사용하도록 구성된 클러스터를 지원하지 않습니다.

다음 각 작업을 수행하려면 먼저 Amazon EMR 클러스터를 시작하고 DynamoDB에서 데이터 위치를 지정한 다음, Hive 명령을 실행하여 DynamoDB에 저장된 데이터를 조작합니다.

Amazon EMR 클러스터를 시작하는 방법에는 여러 가지가 있습니다. Amazon EMR 콘솔, 명령줄 인터페이스(CLI)를 사용하거나 AWS SDK 또는 Amazon EMR API를 사용하여 클러스터를 프로그래밍할 수 있습니다. 그 밖에 Hive 클러스터를 대화식으로 실행할지, 아니면 스크립트에서 실행할지 선택할 수도 있습니다. 이 섹션에서는 Amazon EMR 콘솔과 CLI를 사용하여 대화형 Hive 클러스터를 시작하는 방법에 대해 살펴보겠습니다.

대화식 Hive 사용은 쿼리 성능을 테스트하고 애플리케이션을 조정하는 데 효과적인 방법입니다. 정기적으로 실행할 Hive 명령 세트를 설정한 후 Amazon EMR에서 자동으로 실행할 수 있는 Hive 스크립트를 생성할 수 있습니다.

> **Warning**
>
> DynamoDB 테이블에서 Amazon EMR 읽기 또는 쓰기 작업은 구성된 프로비저닝된 처리량을 소모하여 프로비저닝된 처리량 예외의 발생 주기가 늘어날 가능성이 큽니다. 그래서 Amazon EMR은 대용량 요청을 대비해 지수 백오프 알고리즘을 통한 재시도를 구현함으로써 DynamoDB 테이블에 대한 요청 로드를 관리합니다. Amazon EMR 작업을 다른 트래픽과 동시에 실행하면 할당된 프로비저닝된 처리량 수준을 초과할 수 있습니다. Amazon CloudWatch의 ThrottleRequests 지표를 확인하여 이 상태를 모니터링할 수 있습니다. 요청 로드가 너무 높으

면 클러스터를 다시 시작한 후 [읽기 비율 설정] 또는 [쓰기 비율 설정]을 더 낮은 값으로 설정하여 Amazon EMR 작업을 조절할 수 있습니다. DynamoDB 처리량 설정에 대한 자세한 내용은 [프로비저닝된 처리량]을 참조하세요.

테이블이 [온디맨드 모드]로 구성된 경우 내보내기 또는 가져오기 작업을 실행하기 전에 테이블을 프로비저닝된 모드로 다시 변경해야 합니다. DynamoDB 테이블에서 사용할 리소스를 계산하려면 파이프라인에 처리량 비율이 필요합니다. 온디맨드 모드는 프로비저닝된 처리량을 제거합니다. 처리량 용량을 프로비저닝하려는 경우 Amazon CloudWatch Events 지표를 사용하여 테이블에서 사용하는 총 처리량을 평가할 수 있습니다.

주제

Hive 테이블을 설정하여 Hive 명령 실행

데이터 웨어하우스 애플리케이션인 Apache Hive는 유사 SQL 언어를 사용하여 Amazon EMR 클러스터에 포함된 데이터를 쿼리하는 데 사용할 수 있습니다. Hive에 대한 자세한 내용은 http://hive.apache.org/를 참조하십시오.

다음 절차에서는 이미 클러스터를 생성하고 Amazon EC2 키 페어를 지정했다고 가정합니다. 클러스터 생성을 시작하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 [Amazon EMR 시작하기]를 참조하세요.

MapReduce를 사용하도록 Hive 구성

DynamoDB 테이블을 쿼리하기 위해 Amazon EMR에서 Hive를 사용하는 경우 Hive에 기본 실행 엔진인 Tez가 사용된다면 오류가 발생할 수 있습니다. 이런 이유로 이 섹션에서 설명한 대로 DynamoDB와 통합되는 Hive를 사용해 클러스터를 생성하는 경우 MapReduce를 사용하도록 Hive를 설정하는 구성 분류를 사용하는 것이 좋습니다. 자세한 내용은 [애플리케이션 구성] 단원을 참조하십시오.

다음 조각은 MapReduce를 Hive용 실행 엔진으로 설정하기 위해 사용하는 구성 분류 및 속성을 보여줍니다.

```
[
```

```
            {
                "Classification": "hive-site",
                "Properties": {
                    "hive.execution.engine": "mr"
                }
            }
        ]
```

Hive 명령을 대화식으로 실행하려면

1. 마스터 노드에 연결합니다. 자세한 내용은 Amazon EMR 관리 안내서에서 SSH를 사용하여 프라이머리 노드에 연결을 참조하세요.
2. 현재 마스터 노드에 대한 명령 프롬프트에 `hive`를 입력합니다.

   Hive 프롬프트 `hive>`가 표시됩니다.
3. Hive 애플리케이션의 테이블을 DynamoDB의 데이터에 매핑하는 Hive 명령을 입력합니다. 이 테이블은 Amazon DynamoDB에 저장된 데이터에 대한 참조로 사용되며, 데이터는 Hive에 로컬로 저장되지 않습니다. 따라서 이 테이블을 사용하는 모든 쿼리는 DynamoDB의 실시간 데이터에 대해 실행되므로 명령을 실행할 때마다 테이블의 읽기 또는 쓰기 용량이 소비됩니다. 동일한 데이터 세트에 대해 여러 Hive 명령을 실행하려면 먼저 해당 데이터 세트를 내보내는 것을 고려해 보십시오.

   다음은 Hive 테이블을 DynamoDB 테이블에 매핑하는 구문을 보여 줍니다.

```
CREATE EXTERNAL TABLE hive_tablename
 (hive_column1_name column1_datatype, hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name..."
```

   DynamoDB에서 Hive에 테이블을 생성하는 경우 `EXTERNAL` 키워드를 사용하여 외부 테이블로 생성해야 합니다. 외부 테이블과 내부 테이블의 차이점은 내부 테이블을 삭제하면 내부 테이블의 데이터가 삭제된다는 점입니다. 이는 Amazon DynamoDB에 연결된 경우에 바람직한 동작이 아니므로 외부 테이블만 지원됩니다.

   예를 들어 다음 Hive 명령은 Hive에서 DynamoDB 테이블 dynamodbtable1을 참조하는 hivetable1 테이블을 생성합니다. DynamoDB 테이블 dynamodbtable1에는 해시 및 범위 프라이머리 키 스키

마가 있습니다. 해시 키 요소는 name(문자열 형식)이고, 범위 키 요소는 year(숫자 형식)이고, 각 항목에는 holidays(문자열 세트 형식)에 대한 속성 값이 있습니다.

```
CREATE EXTERNAL TABLE hivetable1 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");
```

라인 1에서는 HiveQL `CREATE EXTERNAL TABLE` 문을 사용합니다. hivetable1에 대해 DynamoDB 테이블의 각 속성 이름 및 값 페어에 대한 열을 설정하고 데이터 형식을 제공해야 합니다. 이러한 값은 대소문자를 구분하지 않으므로 열에 예약어를 제외한 모든 이름을 지정할 수 있습니다.

라인 2에서는 `STORED BY` 문을 사용합니다. `STORED BY`의 값은 Hive와 DynamoDB 사이의 연결을 처리하는 클래스의 이름입니다. `'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'`로 설정되어야 합니다.

행 3에서는 `TBLPROPERTIES` 문을 사용하여 'hivetable1'을 DynamoDB의 올바른 테이블 및 스키마에 연결합니다. `TBLPROPERTIES`에 `dynamodb.table.name` 및 `dynamodb.column.mapping` 파라미터의 값을 제공합니다. 이러한 값은 대소문자를 구분합니다.

Note

테이블에 대한 모든 DynamoDB 속성 이름은 Hive 테이블에 해당 열이 있어야 합니다. Amazon EMR 버전에 따라 다음 시나리오는 일대일 매핑이 존재하지 않는 경우 발생합니다.

- Amazon EMR 5.27.0 이상 버전의 경우, 커넥터에는 DynamoDB 속성 이름과 Hive 테이블에 있는 열 간의 일대일 매핑을 확인하는 검증이 포함되어 있습니다. 일대일 매핑이 존재하지 않으면 오류가 발생합니다.
- Amazon EMR 5.26.0 이하 버전의 경우, Hive 테이블에는 DynamoDB의 이름 및 값 페어가 포함되지 않습니다. DynamoDB 프라이머리 키 속성을 매핑하지 않을 경우 Hive에서 오류가 발생합니다. 비기본 키 속성을 매핑하지 않는 경우에는 오류가 발생하지 않지만

> Hive 테이블에 데이터가 표시되지 않습니다. 데이터 형식이 일치하지 않으면 값은 널입니다.

그런 다음 hivetable1에서 Hive 작업을 실행할 수 있습니다. hivetable1에 대해 실행되는 쿼리는 DynamoDB 계정의 DynamoDB 테이블 dynamodbtable1에 대해 내부적으로 실행되며, 실행할 때마다 읽기 또는 쓰기 유닛을 소비합니다.

DynamoDB 테이블에 대해 Hive 쿼리를 실행하는 경우 충분한 양의 읽기 용량 유닛을 프로비저닝했는지 확인해야 합니다.

예를 들어 DynamoDB 테이블에 읽기 용량 유닛을 100개 프로비저닝하였다고 가정하겠습니다. 그러면 초당 100회의 읽기(또는 409,600바이트)를 수행할 수 있습니다. 테이블에 20GB의 데이터(21,474,836,480바이트)가 포함되어 있고 Hive 쿼리에서 전체 테이블 스캔을 수행하는 경우 쿼리를 실행하는 데 걸리는 시간을 추산할 수 있습니다.

21,474,836,480 / 409,600 = 52,429초 = 14.56시간

필요한 시간을 줄일 수 있는 유일한 방법은 원본 DynamoDB 테이블에서 읽기 용량 유닛을 조정하는 것입니다. Amazon EMR 노드를 더 추가하더라도 효과가 없습니다.

Hive 출력에서 하나 이상의 매퍼 프로세스가 완료되면 진행률이 업데이트됩니다. 하지만 대용량 DynamoDB 테이블에서 프로비저닝된 읽기 용량이 낮게 설정된 경우에는 진행률 출력이 오랫동안 업데이트되지 않을 수도 있습니다. 위 경우에는 작업 진행률이 몇 시간 0%로 표시됩니다. 작업 진행률에 대한 자세한 상태를 보려면 Amazon EMR 콘솔로 이동하여 개별 매퍼 작업 상태와 데이터 읽기에 대한 통계를 볼 수 있습니다. 마스터 노드에서 하둡 인터페이스에 로그온하여 하둡 통계를 볼 수도 있습니다. 여기에는 개별 맵 작업 상태와 일부 데이터 읽기 통계가 표시됩니다. 자세한 정보는 다음의 주제를 참조하세요.

- 프라이머리 노드에 호스팅된 웹 인터페이스
- Hadoop 웹 인터페이스 보기

DynamoDB에서 데이터 내보내기 또는 가져오기 및 테이블 조인과 같은 작업을 수행하는 샘플 HiveQL 문에 대한 자세한 내용은 DynamoDB에서 데이터 내보내기, 가져오기 및 쿼리를 위한 Hive 명령 예제 섹션을 참조하세요.

Hive 요청을 취소하려면

Hive 쿼리를 실행할 때 서버의 초기 응답에는 요청을 취소하는 명령이 포함되어 있습니다. 프로세스 중에 언제든지 요청을 취소하려면 서버 응답에서 Kill Command(Kill 명령)를 사용합니다.

1. 명령줄 클라이언트를 종료하려면 `Ctrl+C`를 입력합니다.
2. 셸 프롬프트에서 요청에 대한 초기 서버 응답의 Kill Command(Kill 명령)를 입력합니다.

   또는 마스터 노드의 명령줄에서 다음 명령을 실행하여 하둡 작업을 종료할 수 있습니다. 여기서 *job-id*는 하둡 작업의 식별자이며 하둡 사용자 인터페이스에서 검색할 수 있습니다.

   ```
   hadoop job -kill job-id
   ```

Hive 및 DynamoDB에 대한 데이터 유형

다음 테이블에는 사용 가능한 Hive 데이터 유형, 해당하는 기본 DynamoDB 유형, 매핑할 수 있는 대체 DynamoDB 유형이 나와 있습니다.

Hive 형식	기본 DynamoDB 유형	대체 DynamoDB 유형
문자열	문자열(S)	
bigint 또는 double	숫자(N)	
이진수	이진수(B)	
boolean	BOOLEAN (BOOL)	
array	list (L)	숫자 세트(NS), 문자열 세트(SS) 또는 이진수 세트(BS)
map<string,string>	item	map (M)
map<string,?>	map (M)	

Hive 형식	기본 DynamoDB 유형	대체 DynamoDB 유형
	null (NULL)	

Hive 데이터를 해당하는 대체 DynamoDB 유형으로 쓰거나 DynamoDB 데이터에 대체 DynamoDB 유형의 대체값이 포함된 경우 `dynamodb.type.mapping` 파라미터를 이용해 열 및 DynamoDB 유형을 지정할 수 있습니다. 다음 예제는 대체 형식 매핑을 지정하기 위한 구문을 보여줍니다.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.type.mapping" = "hive_column1_name:dynamodb_attribute1_datatype");
```

형식 매핑 파라미터는 선택 사항이며 대체 형식을 사용하는 열에 대해서만 지정해야 합니다.

예를 들어 다음 Hive 명령은 DynamoDB 테이블 `dynamodbtable2`를 참조하는 `hivetable2`라는 테이블을 생성합니다. `col3` 열을 문자열 집합(SS) 형식으로 매핑하는 것을 제외하고는 `hivetable1`와 유사합니다.

```
CREATE EXTERNAL TABLE hivetable2 (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable2",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays",
"dynamodb.type.mapping" = "col3:SS");
```

Hive에서 `hivetable1` 및 `hivetable2`는 동일합니다. 하지만 이러한 테이블에 있는 데이터는 해당하는 DynamoDB 테이블에 기록되며, `dynamodbtable1`은 목록을 포함하고 `dynamodbtable2`는 문자열 세트를 포함합니다.

Hive `null` 값을 DynamoDB `null` 유형의 속성으로 작성하려는 경우 `dynamodb.null.serialization` 파라미터를 사용하면 됩니다. 다음 예제는 `null` 직렬화를 지정하기 위한 구문을 보여줍니다.

```
CREATE EXTERNAL TABLE hive_tablename (hive_column1_name column1_datatype,
 hive_column2_name column2_datatype...)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
```

```
TBLPROPERTIES ("dynamodb.table.name" = "dynamodb_tablename",
"dynamodb.column.mapping" =
 "hive_column1_name:dynamodb_attribute1_name,hive_column2_name:dynamodb_attribute2_name...",
"dynamodb.null.serialization" = "true");
```

Null 직렬화 파라미터는 선택 사항이며 지정하지 않으면 `false`로 설정됩니다. DynamoDB `null` 속성은 파라미터 설정에 상관없이 Hive의 `null` 값으로 읽힙니다. `null` 값을 가진 Hive 컬렉션은 null 직렬화 파라미터가 `true`로 설정된 경우에만 DynamoDB에 기록됩니다. 그렇지 않으면 Hive 오류가 발생합니다.

Hive의 bigint 형식은 Java long 형식과 동일하고, Hive double 형식은 정밀도의 측면에서 Java double 형식과 동일합니다. 즉, Hive 데이터 유형에서 사용할 수 있는 것보다 더 높은 정밀도를 가진 숫자 데이터가 DynamoDB에 저장되어 있는 경우 Hive를 사용하여 DynamoDB 데이터를 내보내거나, 가져오거나, 참조하면 정밀도가 손실되거나 Hive 쿼리 오류가 발생할 수 있습니다.

DynamoDB에서 Amazon Simple Storage Service(S3) 또는 HDFS로 내보낸 이진수 유형은 Base64로 인코딩된 문자열로 저장됩니다. 데이터를 Amazon S3 또는 HDFS에서 DynamoDB 이진수 유형으로 가져올 경우 데이터를 Base64 문자열로 인코딩해야 합니다.

Hive 옵션

다음 Hive 옵션을 설정하여 Amazon DynamoDB에서의 데이터 전송을 관리할 수 있습니다. 이러한 옵션은 현재 Hive 세션 동안만 지속됩니다. Hive 명령 프롬프트를 닫았다가 나중에 클러스터에서 다시 열면 해당 설정은 기본값으로 되돌아갑니다.

Hive 옵션	설명
`dynamodb.throughput.read.percent`	DynamoDB 프로비저닝 처리 속도를 테이블에 할당된 범위로 유지하도록 읽기 작업 속도를 설정합니다. 값은 `0.1`과 `1.5` 사이입니다. 값 0.5는 기본 읽기 속도이며, 이 설정에서 Hive는 테이블에서 읽기 프로비저닝된 처리량 리소스의 1/2을 소비하려고 시도합니다. 이 값을 0.5보다 크게 설정하면 읽기 요청 속도가 빨라집니다. 이 값을 0.5보다 작게 설정하면 읽기 요청 속도가 느려집니다. 이 읽기 속도는 근사값입니다. 실제 읽기 속도는 DynamoDB의 키 분포가 균일한지 여부와 같은 요소에 따라 달라집니다.

Hive 옵션	설명
	Hive 작업에서 프로비저닝된 처리량을 자주 초과하거나 실시간 읽기 트래픽 병목 현상이 너무 자주 발생하는 경우 이 값을 `0.5`보다 작게 줄이십시오. 용량이 충분하고 Hive 작업을 더 빠르게 수행하려면 이 값을 `0.5`보다 크게 설정하십시오. 사용되지 않은 입력/출력 작업이 사용 가능한 경우 이 값을 최대 1.5까지 초과 설정할 수도 있습니다.
`dynamodb.throughput.write.percent`	DynamoDB 프로비저닝 처리 속도를 테이블에 할당된 범위로 유지하도록 쓰기 작업 속도를 설정합니다. 값은 `0.1`과 `1.5` 사이입니다. 값 0.5는 기본 쓰기 속도이며, 이 설정에서 Hive는 테이블에서 쓰기 프로비저닝된 처리량 리소스의 1/2을 소비하려고 시도합니다. 이 값을 0.5보다 크게 설정하면 쓰기 요청 속도가 빨라집니다. 이 값을 0.5보다 작게 설정하면 쓰기 요청 속도가 느려집니다. 이 쓰기 속도는 근사값입니다. 실제 쓰기 속도는 DynamoDB의 키 분포가 균일한지 여부와 같은 요소에 따라 달라집니다. Hive 작업에서 프로비저닝된 처리량을 자주 초과하거나 실시간 쓰기 트래픽 병목 현상이 너무 자주 발생하는 경우 이 값을 `0.5`보다 작게 줄이십시오. 용량이 충분하고 Hive 작업을 더 빠르게 수행하려면 이 값을 `0.5`보다 크게 설정하십시오. 사용되지 않은 입력/출력 작업이 사용 가능한 경우나 테이블에 데이터를 처음으로 업로드하고 아직 실시간 트래픽이 없는 경우 이 값을 최대 1.5까지 초과 설정할 수도 있습니다.
`dynamodb.endpoint`	DynamoDB 서비스에 대한 엔드포인트를 지정합니다. 사용 가능한 DynamoDB 엔드포인트에 대한 자세한 내용은 [리전 및 엔드포인트]를 참조하세요.
`dynamodb.max.map.tasks`	DynamoDB에서 데이터를 읽을 때 최대 맵 작업 수를 지정합니다. 이 값은 1보다 크거나 같아야 합니다.

Hive 옵션	설명
`dynamodb.retry.duration`	Hive 명령 재시도를 위한 제한 시간으로 사용할 시간(분)을 지정합니다. 이 값은 0보다 크거나 같은 정수입니다. 기본 제한 시간은 1분입니다.

다음 예제에 표시된 것처럼 이러한 값은 `SET` 명령을 사용하여 설정합니다.

```
SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

DynamoDB에서 데이터 내보내기, 가져오기 및 쿼리를 위한 Hive 명령 예제

다음 예제에서는 Hive 명령을 사용하여 Amazon S3 또는 HDFS로 데이터 내보내기, DynamoDB로 데이터 가져오기, 테이블 조인, 테이블 쿼리 등의 작업을 수행합니다.

Hive 테이블에서의 작업은 DynamoDB에 저장된 데이터를 참조합니다. Hive 명령에는 DynamoDB 테이블의 프로비저닝된 처리량 설정이 적용되며 가져오는 데이터에는 DynamoDB에서 Hive 작업 요청을 처리할 때 DynamoDB 테이블에 쓰여진 데이터가 포함됩니다. 데이터 검색 프로세스가 오래 걸리면 Hive 명령 시작 이후 Hive 명령이 반환하는 일부 데이터가 DynamoDB에서 업데이트된 것일 수 있습니다.

Hive 명령 `DROP TABLE` 및 `CREATE TABLE`은 Hive의 로컬 테이블에만 적용되며 DynamoDB에서 테이블을 생성하거나 삭제하지 않습니다. Hive 쿼리가 DynamoDB의 테이블을 참조할 경우 쿼리를 실행하기 전에 해당 테이블이 이미 있어야 합니다. DynamoDB에서 테이블을 생성하고 삭제하는 방법에 대한 자세한 내용은 Amazon DynamoDB 개발자 안내서에서 DynamoDB의 테이블 작업을 참조하세요.

> **Note**
>
> Hive 테이블을 Amazon S3의 위치에 매핑하는 경우 버킷의 루트 경로인 s3://amzn-s3-demo-bucket에 매핑하지 않습니다. 이 경우 Hive에서 Amazon S3에 데이터를 쓸 때 오류가 발생할 수 있습니다. 대신, 버킷의 하위 경로인 s3://amzn-s3-demo-bucket/mypath에 테이블을 매핑합니다.

DynamoDB에서 데이터 내보내기

Hive를 사용하여 DynamoDB에서 데이터를 내보낼 수 있습니다.

Amazon S3 버킷으로 DynamoDB 테이블을 내보내는 방법

- DynamoDB에 저장된 데이터를 참조하는 Hive 테이블을 생성합니다. 그러면 INSERT OVERWRITE 명령을 호출하여 외부 디렉터리에 데이터를 쓸 수 있습니다. 다음 예제에서 *s3://amzn-s3-demo-bucket/path/subpath/*는 Amazon S3의 유효한 경로입니다. DynamoDB의 값에 맞게 CREATE 명령에서 열과 데이터 형식을 조정합니다. 이 항목을 사용하여 Amazon S3에서 DynamoDB 데이터의 아카이브를 생성할 수 있습니다.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE DIRECTORY 's3://amzn-s3-demo-bucket/path/subpath/' SELECT *
FROM hiveTableName;
```

형식을 사용하여 DynamoDB 테이블을 Amazon S3 버킷으로 내보내는 방법

- Amazon S3 내 위치를 참조하는 외부 테이블을 생성합니다. 이 작업은 s3_export로 아래에 표시됩니다. CREATE 호출 중에 테이블의 행 형식을 지정합니다. 그러면 INSERT OVERWRITE를 사용하여 DynamoDB에서 s3_export로 데이터를 내보낼 때 지정된 형식으로 데이터가 쓰여집니다. 다음 예제에서는 쉼표로 구분된 값(CSV)으로 데이터가 쓰여집니다.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';
```

```
INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

열 매핑을 지정하지 않고 DynamoDB 테이블을 Amazon S3 버킷으로 내보내는 방법

- DynamoDB에 저장된 데이터를 참조하는 Hive 테이블을 생성합니다. 열 매핑을 지정하지 않는다는 점을 제외하면 앞의 예제와 비슷합니다. 테이블에는 `map<string, string>` 형식의 열 1개만 있어야 합니다. Amazon S3에 `EXTERNAL` 테이블을 생성하면 `INSERT OVERWRITE` 명령을 직접 호출하여 DynamoDB에서 에 데이터를 쓸 수 있습니다. 이 항목을 사용하여 Amazon S3에서 DynamoDB 데이터의 아카이브를 생성할 수 있습니다. 열 매핑이 없으므로 이 방법으로 내보낸 테이블은 쿼리할 수 없습니다. Amazon EMR AMI 2.2.x 이상에서 지원되는 Hive 0.8.1.5 이상에서는 열 매핑을 지정하지 않고 데이터를 내보낼 수 있습니다.

```
CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3TableName SELECT *
FROM hiveTableName;
```

데이터 압축을 사용하여 DynamoDB 테이블을 Amazon S3 버킷으로 내보내는 방법

- Hive는 Hive 세션 중에 설정할 수 있는 여러 압축 코덱을 제공합니다. 그러면 내보낸 데이터가 지정된 형식으로 압축됩니다. 다음 예제에서는 Lempel-Ziv-Oberhumer(LZO) 알고리즘을 사용하여 내보낸 파일을 압축합니다.

```
SET hive.exec.compress.output=true;
SET io.seqfile.compression.type=BLOCK;
SET mapred.output.compression.codec = com.hadoop.compression.lzo.LzopCodec;
```

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE lzo_compression_table (line STRING)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE lzo_compression_table SELECT *
FROM hiveTableName;
```

다음과 같은 압축 코덱을 사용할 수 있습니다.

- org.apache.hadoop.io.compress.GzipCodec
- org.apache.hadoop.io.compress.DefaultCodec
- com.hadoop.compression.lzo.LzoCodec
- com.hadoop.compression.lzo.LzopCodec
- org.apache.hadoop.io.compress.BZip2Codec
- org.apache.hadoop.io.compress.SnappyCodec

DynamoDB 테이블을 HDFS로 내보내는 방법

- 다음 Hive 명령을 사용합니다. 여기서 *hdfs:///directoryName*은 유효한 HDFS 경로이고 *hiveTableName*은 DynamoDB를 참조하는 Hive의 테이블입니다. Amazon S3로 데이터를 내보낼 때 Hive 0.7.1.1에서 HDFS를 중간 단계로 사용하므로 이 내보내기 작업은 DynamoDB 테이블을 Amazon S3로 내보내는 것보다 빠릅니다. 또한 다음 예제에서는 `dynamodb.throughput.read.percent`를 1.0으로 설정하여 읽기 요청 빈도를 높이는 방법을 보여 줍니다.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");
```

```
SET dynamodb.throughput.read.percent=1.0;

INSERT OVERWRITE DIRECTORY 'hdfs:///directoryName' SELECT * FROM hiveTableName;
```

또한 Amazon S3로 내보내기 위해 위에 나온 대로 형식 및 압축을 사용하여 HDFS로 데이터를 내보낼 수 있습니다. 그러려면 위 예제의 Amazon S3 디렉터리를 HDFS 디렉터리로 바꾸기만 하면 됩니다.

인쇄할 수 없는 UTF-8 문자 데이터를 Hive로 읽으려면

- 테이블을 생성할 때 STORED AS SEQUENCEFILE 절을 사용하여 인쇄할 수 없는 UTF-8 문자를 Hive로 읽고 쓸 수 있습니다. SequenceFile은 하둡 이진 파일 형식입니다. 이 파일을 읽으려면 하둡을 사용해야 합니다. 다음 예제에서는 DynamoDB에서 Amazon S3로 데이터를 내보내는 방법을 보여 줍니다. 이 기능을 사용하여 인쇄할 수 없는 UTF-8 인코딩 문자를 처리할 수 있습니다.

```
CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");


CREATE EXTERNAL TABLE s3_export(a_col string, b_col bigint, c_col array<string>)
STORED AS SEQUENCEFILE
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

INSERT OVERWRITE TABLE s3_export SELECT *
FROM hiveTableName;
```

DynamoDB로 데이터 가져오기

Hive를 사용하여 DynamoDB에 데이터를 쓸 때 쓰기 용량 유닛 수가 클러스터의 매퍼 수보다 큰지 확인해야 합니다. 예를 들어 m1.xlarge EC2 인스턴스에서 실행되는 클러스터는 인스턴스당 8개의 매퍼를 생성합니다. 클러스터의 인스턴스가 10개일 경우 매퍼는 총 80개입니다. 쓰기 용량 단위가 클러스터의 매퍼 수보다 크지 않으면 Hive 쓰기 작업이 쓰기 처리량을 모두 소비하거나 프로비저닝된 것보다 많은 처리량을 소비하려고 합니다. 각 EC2 인스턴스 유형에서 생성되는 매퍼 수에 대한 자세한 내용은 [하둡 구성](#) 단원을 참조하십시오.

하둡의 매퍼 수는 입력 분할로 제어됩니다. 분할이 너무 적으면 쓰기 명령이 사용 가능한 쓰기 처리량을 모두 사용하지 못할 수도 있습니다.

동일한 키를 가진 항목이 대상 DynamoDB 테이블에 있으면 이 항목을 덮어씁니다. 해당 키를 가진 항목이 대상 DynamoDB 테이블에 없으면 항목이 삽입됩니다.

Amazon S3에서 DynamoDB로 데이터를 가져오는 방법

- Amazon EMR(Amazon EMR) 및 Hive를 사용하여 Amazon S3에서 DynamoDB에 데이터를 쓸 수 있습니다.

```
CREATE EXTERNAL TABLE s3_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM s3_import;
```

열 매핑을 지정하지 않고 Amazon S3 버킷에서 DynamoDB로 테이블을 가져오는 방법

- 전에 DynamoDB에서 내보내어 Amazon S3에 저장된 데이터를 참조하는 EXTERNAL 테이블을 생성합니다. 가져오기 전에 테이블이 DynamoDB에 있으며 전에 내보낸 DynamoDB 테이블과 동일한 키 스키마를 가지고 있는지 확인합니다. 또한 테이블에는 `map<string, string>` 형식의 열 1개만 있어야 합니다. DynamoDB에 연결된 Hive 테이블을 생성하면 `INSERT OVERWRITE` 명령을 호출하여 Amazon S3에서 DynamoDB에 데이터를 쓸 수 있습니다. 열 매핑이 없으므로 이 방법으로 가져온 테이블은 쿼리할 수 없습니다. Amazon EMR AMI 2.2.3 이상에서 지원되는 Hive 0.8.1.5 이상에서는 열 매핑을 지정하지 않고 데이터를 가져올 수 있습니다.

```
CREATE EXTERNAL TABLE s3TableName (item map<string, string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY '\t' LINES TERMINATED BY '\n'
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

CREATE EXTERNAL TABLE hiveTableName (item map<string,string>)
```

```
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1");

INSERT OVERWRITE TABLE hiveTableName SELECT *
FROM s3TableName;
```

HDFS에서 DynamoDB로 테이블을 가져오는 방법

- Amazon EMR 및 Hive를 사용하여 HDFS에서 DynamoDB에 데이터를 쓸 수 있습니다.

```
CREATE EXTERNAL TABLE hdfs_import(a_col string, b_col bigint, c_col array<string>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 'hdfs:///directoryName';

CREATE EXTERNAL TABLE hiveTableName (col1 string, col2 bigint, col3 array<string>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "dynamodbtable1",
"dynamodb.column.mapping" = "col1:name,col2:year,col3:holidays");

INSERT OVERWRITE TABLE hiveTableName SELECT * FROM hdfs_import;
```

DynamoDB의 데이터 쿼리

다음 예제에서는 Amazon EMR을 사용하여 DynamoDB에 저장된 데이터를 쿼리할 수 있는 여러 가지 방법을 보여줍니다.

매핑된 열의 가장 큰 값을 찾으려면(**max**)

- 다음과 같은 Hive 명령을 사용합니다. 첫 번째 명령에서 CREATE 문은 DynamoDB에 저장된 데이터를 참조하는 Hive 테이블을 생성합니다. 그러면 SELECT 문이 해당 테이블을 사용하여 DynamoDB에 저장된 데이터를 쿼리합니다. 다음 예제에서는 제공된 고객이 한 가장 큰 주문을 찾습니다.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
```

```
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT max(total_cost) from hive_purchases where customerId = 717;
```

GROUP BY 절을 사용하여 데이터를 집계하려면

- GROUP BY 절을 사용하여 여러 레코드에서 데이터를 수집할 수 있습니다. 이 절은 종종 sum, count, min 또는 max와 같은 집계 함수와 함께 사용됩니다. 다음 예제에서는 4번 이상 주문한 고객의 가장 큰 주문 목록을 반환합니다.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

SELECT customerId, max(total_cost) from hive_purchases GROUP BY customerId HAVING
 count(*) > 3;
```

두 DynamoDB 테이블을 조인하는 방법

- 다음 예제에서는 Hive 테이블 2개를 DynamoDB에 저장된 데이터에 매핑합니다. 그러면 두 테이블에 조인이 호출됩니다. 조인은 클러스터에서 계산되어 반환됩니다. DynamoDB에서는 조인이 일어나지 않습니다. 이 예제에서는 고객 및 3번 이상 주문한 고객의 구매 목록이 반환합니다.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");
```

```
CREATE EXTERNAL TABLE hive_customers(customerId bigint, customerName string,
 customerAddress array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Customers",
"dynamodb.column.mapping" =
 "customerId:CustomerId,customerName:Name,customerAddress:Address");

Select c.customerId, c.customerName, count(*) as count from hive_customers c
JOIN hive_purchases p ON c.customerId=p.customerId
GROUP BY c.customerId, c.customerName HAVING count > 2;
```

서로 다른 소스의 두 테이블을 조인하려면

- 다음 예제에서 Customer_S3는 Amazon S3에 저장된 CSV 파일을 로드하는 Hive 테이블이며, hive_purchases는 DynamoDB의 데이터를 참조하는 테이블입니다. 다음 예제에서는 Amazon S3에 CSV 파일로 저장된 고객 데이터를 DynamoDB에 저장된 주문 데이터와 조인하여 이름에 'Miller'가 포함된 고객이 한 주문을 나타내는 데이터 세트를 반환합니다.

```
CREATE EXTERNAL TABLE hive_purchases(customerId bigint, total_cost double,
 items_purchased array<String>)
STORED BY 'org.apache.hadoop.hive.dynamodb.DynamoDBStorageHandler'
TBLPROPERTIES ("dynamodb.table.name" = "Purchases",
"dynamodb.column.mapping" =
 "customerId:CustomerId,total_cost:Cost,items_purchased:Items");

CREATE EXTERNAL TABLE Customer_S3(customerId bigint, customerName string,
 customerAddress array<String>)
ROW FORMAT DELIMITED FIELDS TERMINATED BY ','
LOCATION 's3://amzn-s3-demo-bucket/path/subpath/';

Select c.customerId, c.customerName, c.customerAddress from
Customer_S3 c
JOIN hive_purchases p
ON c.customerid=p.customerid
where c.customerName like '%Miller%';
```

> **Note**
>
> 위의 예제에서 명확성과 완성도를 위해 CREATE TABLE 문이 각 예제에 포함되었습니다. 제공된 Hive 테이블에 대해 여러 쿼리나 내보내기 작업을 실행할 때는 Hive 세션을 시작할 때 한 번만 테이블을 생성해야 합니다.

DynamoDB에서 Amazon EMR 작업 성능 최적화

DynamoDB 테이블에서 Amazon EMR 작업은 읽기 작업으로 간주되어 테이블의 프로비저닝 처리량 설정이 적용됩니다. Amazon EMR은 자체 로직을 구현하여 에 따릅니다. 테이블에 대한 로드 밸런스를 맞춤으로써 프로비저닝 처리량을 초과할 가능성을 최소화합니다. 각 Hive 쿼리가 끝날 때마다 Amazon EMR이 프로비저닝 처리량의 초과 횟수를 포함해 쿼리 처리에 사용된 클러스터 정보를 반환합니다. 이 정보는 DynamoDB 처리량에 대한 CloudWatch 지표와 함께 후속 요청에서 DynamoDB 테이블에 대한 로드를 효과적으로 관리하는 데 사용됩니다.

다음은 DynamoDB 테이블 작업 시 Hive 쿼리 성능에 영향을 미치는 요인입니다.

프로비저닝된 읽기 용량 유닛

DynamoDB 테이블에 대해 Hive 쿼리를 실행하는 경우 충분한 양의 읽기 용량 유닛을 프로비저닝했는지 확인해야 합니다.

예를 들어 DynamoDB 테이블에 읽기 용량 단위를 100개 프로비저닝하였다고 가정하겠습니다. 그러면 초당 100회의 읽기(또는 409,600바이트)를 수행할 수 있습니다. 테이블에 20GB의 데이터(21,474,836,480바이트)가 포함되어 있고 Hive 쿼리에서 전체 테이블 스캔을 수행하는 경우 쿼리를 실행하는 데 걸리는 시간을 추산할 수 있습니다.

21,474,836,480 / 409,600 = 52,429초 = 14.56시간

필요한 시간을 줄일 수 있는 유일한 방법은 원본 DynamoDB 테이블에서 읽기 용량 유닛을 조정하는 것입니다. 노드를 Amazon EMR 클러스터에 더 추가하더라도 효과가 없습니다.

Hive 출력에서 하나 이상의 매퍼 프로세스가 완료되면 진행률이 업데이트됩니다. 하지만 대용량 DynamoDB 테이블에서 프로비저닝된 읽기 용량이 낮게 설정된 경우에는 진행률 출력이 오랫동안 업데이트되지 않을 수도 있습니다. 위 경우에는 작업 진행률이 몇 시간 0%로 표시됩니다. 작업 진행률에 대한 자세한 상태를 보려면 Amazon EMR 콘솔로 이동하여 개별 매퍼 작업 상태와 데이터 읽기에 대한 통계를 볼 수 있습니다.

마스터 노드에서 하둡 인터페이스에 로그온하여 하둡 통계를 볼 수도 있습니다. 여기에는 개별 맵 작업 상태와 일부 데이터 읽기 통계가 표시됩니다. 자세한 내용은 Amazon EMR 관리 안내서에서 프라이머리 노드에 호스팅된 웹 인터페이스를 참조하세요.

읽기 비율 설정

기본적으로 Amazon EMR은 현재 프로비저닝된 처리량에 따라 DynamoDB 테이블에 대한 요청 로드를 관리합니다. 하지만 Amazon EMR이 반환하는 작업 정보에 프로비저닝된 처리량을 초과하는 높은 응답 수가 포함된 경우 Hive 테이블을 설정하면서 `dynamodb.throughput.read.percent` 파라미터를 사용해 기본 읽기 비율을 조정할 수 있습니다. 읽기 비율 파라미터 설정에 대한 자세한 내용은 Hive 옵션 단원을 참조하십시오.

쓰기 비율 설정

기본적으로 Amazon EMR은 현재 프로비저닝된 처리량에 따라 DynamoDB 테이블에 대한 요청 로드를 관리합니다. 하지만 Amazon EMR이 반환하는 작업 정보에 프로비저닝된 처리량을 초과하는 높은 응답 수가 포함된 경우 Hive 테이블을 설정하면서 `dynamodb.throughput.write.percent` 파라미터를 사용해 기본 쓰기 비율을 조정할 수 있습니다. 쓰기 비율 파라미터 설정에 대한 자세한 내용은 Hive 옵션 단원을 참조하십시오.

재시도 지속 시간 설정

기본적으로 Amazon EMR은 기본 재시도 주기인 2분 내에 결과를 반환하지 않으면 Hive 쿼리를 다시 실행합니다. 이 주기는 Hive 쿼리 실행 시 `dynamodb.retry.duration` 파라미터를 설정하면 조정 가능합니다. 쓰기 비율 파라미터 설정에 대한 자세한 내용은 Hive 옵션 단원을 참조하십시오.

맵 작업 수

DynamoDB에 저장된 데이터 내보내기 및 쿼리 요청을 처리할 목적으로 Hadoop에서 실행하는 매퍼 대몬(daemon)은 최대 읽기 비율이 초당 1MiB이기 때문에 사용할 수 있는 읽기 용량도 제한됩니다. 이때 DynamoDB에서 사용할 수 있는 프로비저닝된 처리량을 추가하면 매퍼 대몬(daemon) 수를 늘려 Hive 내보내기 및 쿼리 작업 성능을 높일 수 있습니다. 이를 위해 클러스터에 속한 EC2 인스턴스 수를 늘릴 수 있습니다. 또는 각 EC2 인스턴스에서 실행되는 매퍼 대몬(daemon) 수를 늘릴 수 있습니다.

클러스터의 EC2 인스턴스 수를 늘리려면 현재 클러스터를 중단하고 EC2 인스턴스 수를 늘려 재실행하면 됩니다. Amazon EMR 콘솔을 사용해 클러스터를 시작한 경우에는 EC2 인스턴스 구성 대화 상자에서 EC2 인스턴스 수를 지정합니다. 또는 CLI에서 클러스터를 시작한 경우에는 `--num-instances` 옵션을 통해 EC2 인스턴스 수를 지정합니다.

인스턴스에서 실행되는 맵 작업 수는 EC2 인스턴스 유형에 따라 다릅니다. 지원되는 EC2 인스턴스 유형과 각 인스턴스가 제공하는 매퍼 수에 대한 자세한 내용은 작업 구성 단원을 참조하십시오. 여기에서 "작업 구성(Task Configuration)" 단원을 보면 각각 지원되는 구성을 알 수 있습니다.

매퍼 데몬의 수를 늘리는 또 한 가지 방법은 하둡의 `mapreduce.tasktracker.map.tasks.maximum` 구성 파라미터 값을 높이는 것입니다. 이 방법은 EC2 인스턴스 수나 크기를 늘리지 않고 매퍼를 추가함으로써 비용을 절감한다는 점에서 이점이 있습니다. 반면에 이 값을 너무 높게 설정하면 클러스터에 속한 EC2 인스턴스의 메모리가 부족할 수도 있다는 것이 단점입니다. `mapreduce.tasktracker.map.tasks.maximum`을 설정하려면 클러스터를 시작하고 mapred-site 구성 분류의 속성으로 `mapreduce.tasktracker.map.tasks.maximum`의 값을 지정합니다. 방법은 다음 예제와 같습니다. 자세한 내용은 애플리케이션 구성 단원을 참조하십시오.

```
{
    "configurations": [
    {
        "classification": "mapred-site",
        "properties": {
            "mapred.tasktracker.map.tasks.maximum": "10"
        }
    }
    ]
}
```

병렬 데이터 요청

한 명 이상의 사용자 또는 하나 이상의 애플리케이션에서 단일 테이블로 데이터 요청이 다수 이루어지면 읽기 프로비저닝 처리량이 한 번에 소모되어 성능이 느려질 수 있습니다.

프로세스 기간

DynamoDB의 데이터 일관성은 각 노드의 읽기 및 쓰기 작업 순서에 따라 달라집니다. Hive 쿼리가 진행 중일 때는 다른 애플리케이션이 새로운 데이터를 DynamoDB 테이블에 로드하거나 기존 데이터를 변경 또는 삭제하기도 합니다. 이 경우 쿼리 실행 중 데이터 변경 사항은 Hive 쿼리 결과에 반영되지 않습니다.

처리량 초과 주의

DynamoDB에 대해 Hive 쿼리를 실행할 때는 프로비저닝된 처리량을 초과하지 않도록 주의해야 합니다. 그렇지 않으면 애플리케이션이 `DynamoDB::Get`을 직접 호출하는 데 필요한 용량이 고갈되기 때

문입니다. 이러한 용량 고갈을 방지하려면 Amazon CloudWatch의 로그를 점검하고 메트릭을 관찰하여 `DynamoDB::Get` 애플리케이션 호출에 따른 읽기 볼륨 및 병목 현상을 정기적으로 모니터링해야 합니다.

요청 시간

DynamoDB 테이블에 대한 수요가 낮은 시간에 Hive 쿼리가 DynamoDB 테이블에 액세스하도록 일정을 조정하면 성능이 향상됩니다. 예를 들어 애플리케이션 사용자 대부분이 샌프란시스코에 거주한다면 대다수가 잠드는 시간인 새벽 4시(PST)에 DynamoDB 데이터베이스의 레코드 업데이트 없이 일일 데이터를 내보내도록 선택할 수 있습니다.

시간 기반 테이블

데이터가 1일 1개 테이블처럼 시간 기반 DynamoDB 테이블의 형태로 구성되는 경우에는 테이블이 활성 상태가 아닐 때 데이터를 내보낼 수 있습니다. 이 기술은 데이터를 Amazon S3에 지속적으로 백업할 때 사용됩니다.

아카이브된 데이터

DynamoDB에 저장된 데이터에 대해 Hive 쿼리를 여러 차례 실행할 계획이면서 애플리케이션이 데이터 아카이브를 허용하는 경우에는 데이터를 HDFS 또는 Amazon S3로 내보내서 DynamoDB가 아닌 데이터 사본에 대해 Hive 쿼리를 실행할 수 있습니다. 그러면 읽기 작업과 프로비저닝 처리량이 절약됩니다.

Kinesis

Amazon EMR 클러스터는 Hive, Pig, MapReduce, Hadoop Streaming API, Cascading 등 Hadoop 에코시스템에서 익숙한 도구를 사용해 Amazon Kinesis 스트림을 직접 읽고 처리할 수 있습니다. 실행 중인 클러스터에서 Amazon S3, Amazon DynamoDB 및 HDFS의 기존 데이터와 Amazon Kinesis의 실시간 데이터를 조인할 수도 있습니다. 사후 처리 활동을 위해 Amazon EMR에서 Amazon S3 또는 DynamoDB로 데이터를 직접 로드할 수 있습니다. Amazon Kinesis 서비스 주요 기능 및 요금에 대한 자세한 내용은 Amazon Kinesis 페이지를 참조하세요.

Amazon EMR과 Amazon Kinesis 통합으로 어떤 작업을 수행할 수 있나요?

다음과 같은 특정 시나리오는 Amazon EMR과 Amazon Kinesis를 통합할 경우 훨씬 쉬워집니다.

- 스트리밍 로그 분석 - 스트리밍 웹 로그를 분석하여 리전, 브라우저 및 액세스 도메인에 따라 몇 분마다 상위 10개의 오류 유형 목록을 생성할 수 있습니다.

- 고객 참여 - Amazon Kinesis의 클릭스트림 데이터를 DynamoDB 테이블에 저장된 광고 캠페인 정보와 조인하여 특정 웹 사이트에 표시되는 가장 효과적인 광고 카테고리를 식별하는 쿼리를 작성할 수 있습니다.
- 애드혹 대화형 쿼리 - Amazon Kinesis 스트림에서 HDFS로 주기적으로 데이터를 로드하고 빠른 대화형 분석 쿼리를 위한 로컬 Impala 테이블로 데이터를 사용하게 할 수 있습니다.

Amazon Kinesis 스트림의 분석 검사

사용자는 Amazon Kinesis 스트림에 대해 반복이라고 하는 주기적인 배치 분석을 실행할 수 있습니다. Amazon Kinesis 스트림 데이터 레코드는 시퀀스 번호를 사용하여 검색되기 때문에 반복 경계는 Amazon EMR이 DynamoDB 테이블에 저장하는 시작 및 종료 시퀀스 번호로 정의됩니다. 예를 들어 `iteration0`이 끝나면 DynamoDB에 종료 시퀀스 번호를 저장하므로 `iteration1` 작업이 시작되면 스트림에서 후속 데이터를 검색할 수 있습니다. 스트림 데이터의 이 반복 매핑을 검사라고 합니다. 자세한 내용은 Kinesis 커넥터를 참조하세요.

반복을 검사했지만 반복을 처리하는 작업에 실패한 경우 Amazon EMR은 해당 반복의 레코드 재처리를 시도합니다.

검사는 다음을 수행할 수 있는 기능입니다.

- 동일한 스트림 및 논리 이름에서 실행된 이전 쿼리에서 시퀀스 번호가 처리된 후에 데이터 처리를 시작합니다.
- 이전 쿼리에서 처리한 Kinesis의 동일한 데이터 배치를 재처리합니다.

검사를 활성화하려면 스크립트에서 `kinesis.checkpoint.enabled` 파라미터를 `true`로 설정합니다. 또한 다음 파라미터를 구성합니다.

구성 설정	설명
kinesis.checkpoint.metastore.table.name	검사 정보가 저장되는 DynamoDB 테이블 이름
kinesis.checkpoint.metastore.hash.key.name	DynamoDB 테이블의 해시 키 이름
kinesis.checkpoint.metastore.hash.range.name	DynamoDB 테이블의 범위 키 이름
kinesis.checkpoint.logical.name	현재 처리의 논리적 이름

구성 설정	설명
kinesis.checkpoint.iteration.no	논리적 이름과 연관된 처리의 반복 번호
kinesis.rerun.iteration.without.wait	실패한 반복이 시간 초과를 기다리지 않고 다시 실행할 수 있는지 여부를 나타내는 부울 값입니다. 기본값은 `false`입니다.

Amazon DynamoDB 테이블의 프로비저닝된 IOPS 권장 사항

Amazon Kinesis용 Amazon EMR 커넥터는 메타데이터 검사를 위한 백업으로 DynamoDB 데이터베이스를 사용합니다. 검사 간격으로 Amazon EMR 클러스터가 포함된 Amazon Kinesis 스트림의 데이터를 사용하기 전에 DynamoDB에서 테이블을 생성해야 합니다. 테이블은 Amazon EMR 클러스터와 같은 리전에 있어야 합니다. 다음은 DynamoDB 테이블에 대해 프로비저닝해야 하는 IOPS 수에 대한 일반적인 권장사항입니다. `j`는 동시에 실행할 수 있는 Hadoop 작업의 최대 수(다른 논리적 이름+반복 수 조합 포함)이고 `s`는 작업이 처리하는 샤드의 최대 수입니다.

읽기 용량 유닛: `j`*s/5

쓰기 용량 유닛: `j`*s

성능 고려 사항

Amazon Kinesis 샤드 처리량은 Amazon EMR 클러스터를 구성하는 노드의 인스턴스 크기 및 스트림의 레코드 크기에 정비례합니다. 프라이머리 및 코어 노드에는 m5.xlarge 또는 더 큰 인스턴스를 사용하는 것이 좋습니다.

Amazon EMR을 사용하여 Amazon Kinesis 분석 예약

제한 시간과 최대 반복 기간 제한이 있는 활성 Amazon Kinesis 스트림에서 데이터를 분석하는 경우, 분석을 자주 실행하여 스트림에서 정기적으로 세부 정보를 수집해야 합니다. 이러한 스크립트와 쿼리를 주기적으로 실행하는 방법에는 여러 가지가 있지만 이처럼 반복적인 작업에는 AWS Data Pipeline을 사용하는 것이 좋습니다. 자세한 내용은 AWS Data Pipeline 개발자 안내서에서 AWS Data Pipeline PigActivity 및 AWS Data Pipeline HiveActivity를 참조하세요.

Spark Kinesis 커넥터를 Amazon EMR 7.0용 SDK 2.x로 마이그레이션

AWS SDK는 자격 증명 관리, S3 및 Kinesis 서비스 연결과 같은 클라우드 컴퓨팅 서비스와 상호 작용할 수 있는 AWS 풍부한 APIs 및 라이브러리 세트를 제공합니다. Spark Kinesis 커넥터는 Kinesis Data

Streams의 데이터를 소비하는 데 사용되며 수신된 데이터는 Spark의 실행 엔진에서 변환 및 처리됩니다. 현재이 커넥터는 1.x SDK AWS 및 Kinesis-client-library(KCL)를 기반으로 구축되었습니다.

AWS SDK 2.x 마이그레이션의 일환으로 Spark Kinesis 커넥터도 SDK 2.x로 실행되도록 적절히 업데이트됩니다. Amazon EMR 7.0 릴리스의 Spark에는 Apache Spark의 커뮤니티 버전에서 아직 사용할 수 없는 SDK 2.x 업그레이드가 포함되어 있습니다. 7.0 미만의 릴리스에서 Spark Kinesis 커넥터를 사용하는 경우 Amazon EMR 7.0으로 마이그레이션하기 전에 SDK 2.x에서 실행되도록 애플리케이션 코드를 마이그레이션해야 합니다.

마이그레이션 안내서

이 섹션에서는 업그레이드된 Spark Kinesis 커넥터로 애플리케이션을 마이그레이션하는 단계를 설명합니다. 여기에는 AWS SDK 2.x의 Kinesis Client Library(KCL) 2.x, AWS 자격 증명 공급자 및 AWS 서비스 클라이언트로 마이그레이션하는 가이드가 포함되어 있습니다. 참고로 Kinesis 커넥터를 사용하는 샘플 WordCount 프로그램도 포함되어 있습니다.

주제

1.x에서 2.x로 KCL 마이그레이션

- **`KinesisInputDStream`의 지표 수준 및 차원**

  `KinesisInputDStream`을 인스턴스화할 때 스트림의 지표 수준 및 차원을 제어할 수 있습니다. 다음 예제는 KCL 1.x로 이러한 파라미터를 사용자 지정하는 방법을 보여줍니다.

```
import
 com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel

val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
```

```
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)

 .metricsEnabledDimensions(KinesisClientLibConfiguration.DEFAULT_METRICS_ENABLED_DIMENSIONS.a
  .build()
```

KCL 2.x에서는 이러한 구성 설정의 패키지 이름이 다릅니다. 2.x로 마이그레이션:

1. 가져오기 명령문을 각각 `com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibCon` 및 `com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel`에서 `software.amazon.kinesis.metrics.MetricsLevel` 및 `software.amazon.kinesis.metrics.MetricsUtil`로 변경하세요.

   ```
   // import com.amazonaws.services.kinesis.metrics.interfaces.MetricsLevel
   import software.amazon.kinesis.metrics.MetricsLevel

   // import
    com.amazonaws.services.kinesis.clientlibrary.lib.worker.KinesisClientLibConfiguration
   import software.amazon.kinesis.metrics.MetricsUtil
   ```

2. `metricsEnabledDimensionsKinesisClientLibConfiguration.DEFAULT_METRICS_ENABL` 줄을 `metricsEnabledDimensionsSet(MetricsUtil.OPERATION_DIMENSION_NAME, MetricsUtil.SHARD_ID_DIMENSION_NAME)`로 바꿉니다.

다음은 지표 수준 및 지표 차원을 사용자 지정한 `KinesisInputDStream`의 업데이트된 버전입니다.

```
import software.amazon.kinesis.metrics.MetricsLevel
import software.amazon.kinesis.metrics.MetricsUtil

val kinesisStream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
```

```
  .checkpointInterval(kinesisCheckpointInterval)
  .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
  .metricsLevel(MetricsLevel.DETAILED)
  .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
 MetricsUtil.SHARD_ID_DIMENSION_NAME))
  .build()
```

- `KinesisInputDStream`의 메시지 핸들러 함수

  파티션 키와 같이 레코드에 포함된 다른 데이터를 사용하려는 경우 `KinesisInputDStream`를 인스턴스화할 때 Kinesis Record를 사용하여 일반 객체 T를 반환하는 "메시지 핸들러 함수"를 제공할 수도 있습니다.

  KCL 1.x에서 메시지 핸들러 함수 서명은 `Record => T`이고. 여기서 Record는 `com.amazonaws.services.kinesis.model.Record`입니다. KCL 2.x에서는 핸들러의 서명이 `KinesisClientRecord => T`로 변경됩니다. 여기서 KinesisClientRecord는 `software.amazon.kinesis.retrieval.KinesisClientRecord`입니다.

  다음은 KCL 1.x에서 메시지 핸들러를 제공하는 예제입니다.

```
import com.amazonaws.services.kinesis.model.Record


def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

  메시지 핸들러를 마이그레이션하는 방법:

  1. `com.amazonaws.services.kinesis.model.Record`의 가져오기 명령문을 `software.amazon.kinesis.retrieval.KinesisClientRecord`로 변경하세요.

```
// import com.amazonaws.services.kinesis.model.Record
```

```
import software.amazon.kinesis.retrieval.KinesisClientRecord
```

2. 메시지 핸들러의 메서드 서명을 업데이트하세요.

```
//def addFive(r: Record): Int = JavaUtils.bytesToString(r.getData).toInt + 5
def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt
 + 5
```

다음은 KCL 2.x에서 메시지 핸들러를 제공하는 업데이트된 예입니다.

```
import software.amazon.kinesis.retrieval.KinesisClientRecord


def addFive = (r: KinesisClientRecord) => JavaUtils.bytesToString(r.data()).toInt + 5
val stream = KinesisInputDStream.builder
  .streamingContext(ssc)
  .streamName(streamName)
  .endpointUrl(endpointUrl)
  .regionName(regionName)
  .initialPosition(new Latest())
  .checkpointAppName(appName)
  .checkpointInterval(Seconds(10))
  .storageLevel(StorageLevel.MEMORY_ONLY)
  .buildWithMessageHandler(addFive)
```

KCL 1.x에서 2.x로 마이그레이션에 대한 자세한 내용은 KCL 1.x에서 KCL 2.x로 소비자 마이그레이션을 참조하세요.

AWS SDK 1.x에서 2.x로 AWS 자격 증명 공급자 마이그레이션

자격 증명 공급자는와 상호 작용하기 위한 AWS 자격 증명을 얻는 데 사용됩니다 AWS. SDK 2.x의 보안 인증 공급자와 관련된 몇 가지 인터페이스 및 클래스 변경 사항은 여기에서 확인할 수 있습니다. Spark Kinesis 커넥터는 1.x 버전의 AWS 자격 증명 공급자를 반환하는 인터페이스 (`org.apache.spark.streaming.kinesis.SparkAWSCredentials`) 및 구현 클래스를 정의했습니다. Kinesis 클라이언트를 초기화할 때 이러한 보안 인증 공급자가 필요합니다. 예를 들어 `SparkAWSCredentials.provider` 애플리케이션에서 메서드를 사용하는 경우 2.x 버전의 AWS 자격 증명 공급자를 사용하도록 코드를 업데이트해야 합니다.

다음은 AWS SDK 1.x에서 자격 증명 공급자를 사용하는 예입니다.

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import com.amazonaws.auth.AWSCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
    .basicCredentials("accessKey", "secretKey")
    .build()

val credentialProvider = basicSparkCredentials.provider
assert(credentialProvider.isInstanceOf[AWSCredentialsProvider], "Type should be
 AWSCredentialsProvider")
```

SDK 2.x로 마이그레이션하는 방법:

1. `com.amazonaws.auth.AWSCredentialsProvider`의 가져오기 명령문을 `software.amazon.awssdk.auth.credentials.AwsCredentialsProvider`로 변경

```
//import com.amazonaws.auth.AWSCredentialsProvider
import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider
```

2. 이 클래스를 사용하는 나머지 코드를 업데이트하세요.

```
import org.apache.spark.streaming.kinesis.SparkAWSCredentials
import software.amazon.awssdk.auth.credentials.AwsCredentialsProvider

val basicSparkCredentials = SparkAWSCredentials.builder
    .basicCredentials("accessKey", "secretKey")
    .build()

val credentialProvider = basicSparkCredentials.provider
assert (credentialProvider.isInstanceOf[AwsCredentialsProvider], "Type should be
 AwsCredentialsProvider")
```

AWS SDK 1.x에서 2.x로 AWS 서비스 클라이언트 마이그레이션

AWS 서비스 클라이언트는 2.x(예: `software.amazon.awssdk`)에서 패키지 이름이 다른 반면, SDK 1.x는를 사용합니다`com.amazonaws`. 클라이언트 변경에 대한 자세한 내용은 여기를 참조하세요. 코드에서 이러한 서비스 클라이언트를 사용하는 경우 그에 따라 클라이언트를 마이그레이션해야 합니다.

다음은 SDK 1.x에서 클라이언트를 생성하는 예제입니다.

```
import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
import com.amazonaws.services.dynamodbv2.document.DynamoDB

AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();
```

2.x로 마이그레이션:

1. 서비스 클라이언트의 가져오기 명령문을 변경하세요. DynamoDB 클라이언트를 예로 들어 보겠습니다. `com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient` 또는 `com.amazonaws.services.dynamodbv2.document.DynamoDB`를 `software.amazon.awssdk.services.dynamodb.DynamoDbClient`로 변경해야 할 것입니다.

   ```
   // import com.amazonaws.services.dynamodbv2.AmazonDynamoDBClient
   // import com.amazonaws.services.dynamodbv2.document.DynamoDB
   import software.amazon.awssdk.services.dynamodb.DynamoDbClient
   ```

2. 클라이언트를 초기화하는 코드를 업데이트하세요.

   ```
   // AmazonDynamoDB ddbClient = AmazonDynamoDBClientBuilder.defaultClient();
   // AmazonDynamoDBClient ddbClient = new AmazonDynamoDBClient();

   DynamoDbClient ddbClient = DynamoDbClient.create();
   DynamoDbClient ddbClient = DynamoDbClient.builder().build();
   ```

   AWS SDK를 1.x에서 2.x로 마이그레이션하는 방법에 대한 자세한 내용은 AWS SDK for Java 1.x와 2.x의 차이점을 참조하세요.

스트리밍 애플리케이션의 코드 예제

```
import java.net.URI
import software.amazon.awssdk.auth.credentials.DefaultCredentialsProvider
import software.amazon.awssdk.http.apache.ApacheHttpClient
import software.amazon.awssdk.services.kinesis.KinesisClient
import software.amazon.awssdk.services.kinesis.model.DescribeStreamRequest
import software.amazon.awssdk.regions.Region
import software.amazon.kinesis.metrics.{MetricsLevel, MetricsUtil}

import org.apache.spark.SparkConf
```

```
import org.apache.spark.storage.StorageLevel
import org.apache.spark.streaming.{Milliseconds, StreamingContext}
import org.apache.spark.streaming.dstream.DStream.toPairDStreamFunctions
import org.apache.spark.streaming.kinesis.KinesisInitialPositions.Latest
import org.apache.spark.streaming.kinesis.KinesisInputDStream


object KinesisWordCountASLSDKV2 {

  def main(args: Array[String]): Unit = {
    val appName = "demo-app"
    val streamName = "demo-kinesis-test"
    val endpointUrl = "https://kinesis.us-west-2.amazonaws.com"
    val regionName = "us-west-2"

    // Determine the number of shards from the stream using the low-level Kinesis
 Client
    // from the AWS Java SDK.
    val credentialsProvider = DefaultCredentialsProvider.create
    require(credentialsProvider.resolveCredentials() != null,
      "No AWS credentials found. Please specify credentials using one of the methods
 specified " +
        "in https://docs.aws.amazon.com/sdk-for-java/latest/developer-guide/
credentials.html")
    val kinesisClient = KinesisClient.builder()
      .credentialsProvider(credentialsProvider)
      .region(Region.US_WEST_2)
      .endpointOverride(URI.create(endpointUrl))
      .httpClientBuilder(ApacheHttpClient.builder())
      .build()
    val describeStreamRequest = DescribeStreamRequest.builder()
      .streamName(streamName)
      .build()
    val numShards = kinesisClient.describeStream(describeStreamRequest)
      .streamDescription
      .shards
      .size


    // In this example, we are going to create 1 Kinesis Receiver/input DStream for
 each shard.
    // This is not a necessity; if there are less receivers/DStreams than the number of
 shards,
```

```
    // then the shards will be automatically distributed among the receivers and each
receiver
    // will receive data from multiple shards.
    val numStreams = numShards

    // Spark Streaming batch interval
    val batchInterval = Milliseconds(2000)

    // Kinesis checkpoint interval is the interval at which the DynamoDB is updated
with information
    // on sequence number of records that have been received. Same as batchInterval for
this
    // example.
    val kinesisCheckpointInterval = batchInterval

    // Setup the SparkConfig and StreamingContext
    val sparkConfig = new SparkConf().setAppName("KinesisWordCountASLSDKV2")
    val ssc = new StreamingContext(sparkConfig, batchInterval)

    // Create the Kinesis DStreams
    val kinesisStreams = (0 until numStreams).map { i =>
      KinesisInputDStream.builder
        .streamingContext(ssc)
        .streamName(streamName)
        .endpointUrl(endpointUrl)
        .regionName(regionName)
        .initialPosition(new Latest())
        .checkpointAppName(appName)
        .checkpointInterval(kinesisCheckpointInterval)
        .storageLevel(StorageLevel.MEMORY_AND_DISK_2)
        .metricsLevel(MetricsLevel.DETAILED)
        .metricsEnabledDimensions(Set(MetricsUtil.OPERATION_DIMENSION_NAME,
MetricsUtil.SHARD_ID_DIMENSION_NAME))
        .build()
    }

    // Union all the streams
    val unionStreams = ssc.union(kinesisStreams)

    // Convert each line of Array[Byte] to String, and split into words
    val words = unionStreams.flatMap(byteArray => new String(byteArray).split(" "))

    // Map each word to a (word, 1) tuple so we can reduce by key to count the words
    val wordCounts = words.map(word => (word, 1)).reduceByKey(_ + _)
```

```
    // Print the first 10 wordCounts
    wordCounts.print()

    // Start the streaming context and await termination
    ssc.start()
    ssc.awaitTermination()
  }
}
```

업그레이드된 Spark Kinesis 커넥터 사용 시 고려 사항

- 애플리케이션에서 JDK 버전 11 미만의 `Kinesis-producer-library`를 사용하는 경우 `java.lang.NoClassDefFoundError: javax/xml/bind/DatatypeConverter`와 같은 예외가 발생할 수 있습니다. 이는 EMR 7.0이 기본적으로 JDK 17과 함께 제공되고 Java 11 이상부터 J2EE 모듈이 표준 라이브러리에서 제거되었기 때문에 발생합니다. 이 문제는 pom 파일에 다음 종속성을 추가하여 해결할 수 있습니다. 상황에 맞게 라이브러리 버전을 교체합니다.

```
<dependency>
      <groupId>javax.xml.bind</groupId>
      <artifactId>jaxb-api</artifactId>
      <version>${jaxb-api.version}</version>
    </dependency>
```

- Spark Kinesis 커넥터 jar는 EMR 클러스터가 생성된 후 `/usr/lib/spark/connector/lib/` 경로에서 찾을 수 있습니다.

S3DistCp(s3-dist-cp)

Apache DistCp는 대량의 데이터를 복사하는 데 사용할 수 있는 오픈 소스 도구입니다. S3DistCp는 DistCp와 비슷하지만 AWS, 특히 Amazon S3에서 함께 작동하도록 최적화되었습니다. Amazon EMR 버전 4.0 및 이후에서 S3DistCp 명령은 `s3-dist-cp`이며, 클러스터 단계로 혹은 명령줄에 추가됩니다. S3DistCp를 사용하여 Amazon S3의 대량 데이터를 HDFS로 효율적으로 복사할 수 있습니다. HDFS에서는 Amazon EMR 클러스터의 후속 단계를 통해 데이터를 처리할 수 있습니다. 또한 S3DistCp를 사용하여 Amazon S3 버킷 간에 또는 HDFS에서 Amazon S3로 데이터를 복사할 수도 있습니다. S3DistCp는 버킷과 AWS 계정 간에 많은 수의 객체를 병렬로 복사하는 데 더 확장 가능하고 효율적입니다.

실제 시나리오에서 S3DistCP의 유연성을 보여주는 특정 명령은 AWS 빅 데이터 블로그의 [S3DistCp 사용에 대한 7가지 팁을] 참조하세요.

DistCp처럼 S3DistCp도 MapReduce를 사용하여 분산 방식으로 복사합니다. 여러 서버 간에 복사, 오류 처리, 복구 및 보고 작업을 공유합니다. Apache DistCp 오픈 소스 프로젝트에 대한 자세한 내용은 Apache Hadoop 설명서에서 DistCp guide를 참조하세요.

S3DistCp가 지정된 파일의 일부 또는 전체를 복사할 수 없는 경우 클러스터 단계가 실패하고 0 이외의 오류 코드를 반환합니다. 이러한 상황이 발생할 경우 S3DistCp는 부분적으로 복사된 파일을 지우지 않습니다.

> **Important**
>
> S3DistCp는 밑줄 문자를 포함하는 Amazon S3 버킷 이름을 지원하지 않습니다.
> S3DistCp는 Parquet 파일의 연결을 지원하지 않습니다. 대신 PySpark를 사용하십시오. 자세한 내용은 Amazon EMR에서 parquet 파일 연결을 참조하세요.
> S3DistCp를 사용하여 단일 파일(디렉터리 대신)을 S3에서 HDFS로 복사할 때 복사 오류를 방지하려면 Amazon EMR 버전 5.33.0 이상 또는 Amazon EMR 버전 6.3.0 이상을 사용합니다.

S3DistCp 옵션

S3DistCp는 DistCp와 비슷하지만 데이터를 복사하고 압축하는 방법을 변경하는 다양한 옵션 세트를 지원합니다.

S3DistCp를 직접 호출하면 다음 테이블에서 설명하는 옵션을 지정할 수 있습니다. 이 옵션은 인수 목록을 사용하여 단계에 추가됩니다. 다음 표에는 S3DistCp 인수 예가 나와 있습니다.

옵션	설명	필수
`--src=LOCATION`	복사할 데이터의 위치입니다. 이 위치는 HDFS 또는 Amazon S3 위치일 수 있습니다. 예시: `--src=s3://amzn-s3-demo-bucket/logs/j-3GYXXXXXX9IOJ/node` **Important** S3DistCp는 밑줄 문자를 포함하는 Amazon S3 버킷 이름을 지원하지 않습니다.	예

옵션	설명	필수
`--dest=LOCATION`	데이터의 대상입니다. 이 위치는 HDFS 또는 Amazon S3 위치일 수 있습니다. 예시: `--dest=hdfs:///output` **Important** S3DistCp는 밑줄 문자를 포함하는 Amazon S3 버킷 이름을 지원하지 않습니다.	예
`--srcPattern=PATTERN`	`--src`에서 데이터의 하위 세트로 복사 작업을 필터링하는 [정규식](https://en.wikipedia.org/wiki/Regular_expression)입니다. `--srcPattern` 과 `--groupBy` 를 둘 다 지정하지 않으면 `--src`의 모든 데이터가 `--dest`로 복사됩니다. 정규식 인수에 별표(*) 같은 특수 문자가 포함되어 있는 경우 정규식 또는 전체 `--args` 문자열을 작은따옴표(')에 포함해야 합니다. 예시: `--srcPattern=.*daemons.*-hadoop-.*`	아니요

옵션	설명	필수
`--groupBy=PATTERN`	S3DistCp에서 식과 일치하는 파일을 연결하는 정규식입니다. 예를 들면 이 옵션을 사용하여 한 시간 동안 작성된 모든 로그 파일을 단일 파일로 결합할 수 있습니다. 연결된 파일 이름은 그룹화 정규식과 일치하는 값입니다. 괄호는 파일을 그룹화하는 방식을 나타내며, 괄호 안에 지정된 문과 일치하는 모든 항목이 단일 출력 파일로 결합됩니다. 괄호에 지정된 문이 정규식에 포함되지 않는 경우 S3DistCp 단계에 대해 클러스터가 실패하며 오류가 발생합니다. 정규식 인수에 별표(*) 같은 특수 문자가 포함되어 있는 경우 정규식 또는 전체 `--args` 문자열을 작은따옴표(')에 포함해야 합니다. `--groupBy` 를 지정하면 지정한 패턴과 일치하는 파일만 복사됩니다. `--groupBy` 와 `--srcPattern` 을 동시에 지정할 필요는 없습니다. 예시: `--groupBy=.*subnetid.*([0-9]+-[0-9]+-[0-9]+-[0-9]+).*`	아니요

옵션	설명	필수
`--targetSize=SIZE`	`--groupBy` 옵션에 따라 생성할 파일의 크기로, 단위는 메비바이트(MiB)입니다. 이 값은 정수여야 합니다. `--targetSize` 를 설정하는 경우 S3DistCp에서는 이 크기에 맞추려고 합니다. 복사한 파일의 실제 크기는 이 값보다 크거나 작을 수도 있습니다. 작업이 데이터 파일의 크기에 따라 집계되므로 대상 파일 크기가 소스 데이터 파일 크기와 일치할 수 있습니다. `--groupBy` 에 의해 연결된 파일이 `--targetSize` 의 값보다 크면 부분 파일로 나눠지고 이름 끝에 숫자 값이 순차적으로 지정됩니다. 예를 들어 `myfile.gz` 에 연결된 파일은 `myfile0.gz` , `myfile1.gz` 등 여러 부분으로 나눠집니다. 예시: `--targetSize=2`	아니요
`--appendToLastFile`	이미 있는 파일을 Amazon S3에서 HDFS로 복사하는 경우 S3DistCp의 동작을 지정합니다. 이 옵션은 새 파일 데이터를 기존 파일에 추가합니다. `--appendToLastFile` 을 `--groupBy` 와 함께 사용할 경우 동일 그룹과 일치하는 파일에 새 데이터가 추가됩니다. 이 옵션은 `--targetSize` 와 함께 사용하는 경우 `--groupBy.` 동작과 관련이 있습니다.	아니요
`--outputCodec=CODEC`	복사한 파일에 사용할 압축 코덱을 지정합니다. 이때 `gzip`, `gz`, `lzo`, `snappy` 또는 `none` 값을 사용할 수 있습니다. 예를 들어 이 옵션을 사용하여 Gzip으로 압축된 입력 파일을 LZO로 압축된 출력 파일로 변환하거나 복사 작업의 일부로서 파일의 압축을 풀 수 있습니다. 출력 코덱을 선택하는 경우 파일 이름에 해당 확장명(예: gz 및 `gzip`의 경우, 확장명은 `.gz`)이 추가됩니다. `--outputCodec` 의 값을 지정하지 않으면 파일이 압축 변화 없이 복사됩니다. 예시: `--outputCodec=lzo`	아니요

옵션	설명	필수
`--s3ServerSideEncryption`	대상 데이터가 SSL을 사용하여 전송되고 AWS 서비스 측 키를 사용하여 Amazon S3에서 자동으로 암호화되는지 확인합니다. S3DistCp를 사용하여 데이터를 검색할 때 객체가 자동으로 암호화 해제됩니다. 암호화되지 않은 객체를 암호화가 요구되는 Amazon S3 버킷으로 복사하려는 경우 작업이 실패합니다. 자세한 내용은 [데이터 암호화 사용](#)을 참조하세요. 예시: `--s3ServerSideEncryption`	아니요
`--deleteOnSuccess`	이 옵션을 지정한 경우 복사 작업이 성공적으로 수행되면 S3DistCp가 복사한 파일이 소스 위치에서 삭제됩니다. 이 옵션은 예약 작업으로서 로그 파일 같은 출력 파일을 한 위치에서 다른 위치로 복사하려 하며 동일한 파일을 두 번 복사하기를 원치 않는 경우에 유용합니다. 예시: `--deleteOnSuccess`	아니요
`--disableMultipartUpload`	멀티파트 업로드 사용을 비활성화합니다. 예시: `--disableMultipartUpload`	아니요
`--multipartUploadChunkSize=SIZE`	Amazon S3 멀티파트 업로드의 각 파트 크기(MiB). S3DistCp는 `multipartUploadChunkSize` 보다 큰 데이터를 복사할 때 멀티파트 업로드를 사용합니다. 작업 성능을 향상시키기 위해 각 부품 크기를 늘릴 수 있습니다. 기본 크기는 128MiB입니다. 예시: `--multipartUploadChunkSize=1000`	아니요
`--numberFiles`	출력 파일 앞에 순차 번호를 붙입니다. `--startingIndex` 에 다른 값을 지정하는 경우가 아니면 번호는 0부터 시작합니다. 예시: `--numberFiles`	아니요

옵션	설명	필수
`--startingIndex=INDEX`	`--numberFiles` 와 함께 사용되어 순서에서 첫 번째 번호를 지정합니다. 예시: `--startingIndex=1`	아니요
`--outputManifest=FILENAME`	S3DistCp에서 복사한 모든 파일의 목록을 포함하는 텍스트 파일(Gzip으로 압축됨)을 생성합니다. 예시: `--outputManifest=manifest-1.gz`	아니요
`--previousManifest=PATH`	`--outputManifest` 플래그를 사용하여 이전 S3DistCp 호출 중에 생성된 매니페스트 파일을 읽습니다. `--previousManifest` 플래그를 설정한 경우 S3DistCp가 매니페스트에 나열된 파일을 복사 작업에서 제외시킵니다. `--outputManifest` 를 `--previousManifest` 와 함께 지정하는 경우 복사되지 않았더라도 이전 매니페스트에 나열된 파일이 새 매니페스트 파일에도 나타납니다. 예시: `--previousManifest=/usr/bin/manifest-1.gz`	아니요
`--requirePreviousManifest`	이전 S3DistCp 호출 중에 생성된 이전 매니페스트가 필요합니다. false로 설정된 경우 이전 매니페스트가 지정되지 않았을 때 오류가 발생하지 않습니다. 기본값은 true입니다.	아니요
`--copyFromManifest`	S3DistCp에서 파일 목록을 복사에서 제외하는 것이 아니라 복사할 파일 목록으로서 지정한 매니페스트 파일을 사용하도록 `--previousManifest` 의 동작을 반대로 설정합니다. 예시: `--copyFromManifest --previousManifest=/usr/bin/manifest-1.gz`	아니요

옵션	설명	필수
--s3Endpoint=ENDPOINT	파일 업로드 시 사용할 Amazon S3 엔드포인트를 지정합니다. 이 옵션은 소스 및 대상 모두의 엔드포인트를 설정합니다. 설정하지 않은 경우 기본 엔드포인트는 s3.amazonaws.com 입니다. Amazon S3 엔드포인트 목록은 리전 및 엔드포인트를 참조하세요. 예시: --s3Endpoint=s3.eu-west-1.amazonaws.com	아니요
--storageClass=CLASS	대상이 Amazon S3일 때 사용할 스토리지 클래스입니다. 유효한 값은 STANDARD 및 REDUCED_REDUNDANCY입니다. 이 옵션을 지정하지 않은 경우 S3DistCp에서 스토리지 클래스를 보존합니다. 예시: --storageClass=STANDARD	아니요
--srcPrefixesFile=PATH	Amazon S3(s3://), HDFS(hdfs:///) 또는 로컬 파일 시스템(file:/)에 있는 src 접두사 목록(줄당 접두사 한 개)을 포함하는 텍스트 파일입니다. srcPrefixesFile 을 제공하는 경우 S3DistCp에서 src 경로를 나열하지 않습니다. 대신에 결합된 결과로서 소스 목록을 생성하여 이 파일에 지정된 모든 접두사를 나열합니다. 이러한 접두사를 사용하는 src 경로와는 달리, 상대 경로를 사용하여 대상 경로를 생성합니다. srcPattern 도 지정한 경우 이 패턴이 소스 접두사의 결합된 목록 결과에 적용되어 입력을 추가로 필터링합니다. copyFromManifest 를 사용하는 경우 매니페스트의 객체가 복사되고 srcPrefixesFile 이 무시됩니다. 예시: --srcPrefixesFile=PATH	아니요

위의 옵션 외에도, S3DistCp는 도구 인터페이스를 구현하므로 일반 옵션이 지원됩니다.

S3DistCp를 클러스터에 단계로 추가

클러스터에서 단계로 추가하여 S3DistCp를 호출할 수 있습니다. 콘솔, CLI 또는 API를 사용하여 시작 시 클러스터에 또는 실행 중인 클러스터에 단계를 추가할 수 있습니다. 다음 예제에서는 S3DistCp를 사용하여 실행 중인 클러스터에 추가하는 방법을 보여 줍니다. 클러스터에 단계를 추가하는 방법에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 클러스터에 작업 제출을 참조하세요.

를 사용하여 실행 중인 클러스터에 S3DistCp 단계를 추가하려면 AWS CLI

에서 Amazon EMR 명령을 사용하는 방법에 대한 자세한 AWS CLI내용은 AWS CLI 명령 참조를 참조하세요.

- S3DistCp를 호출하는 클러스터에 단계를 추가하려면 S3DistCp에서 복사 작업을 수행하는 방식을 지정하는 파라미터를 인수로 전달합니다.

  다음 예제에서는 Amazon S3에서 `hdfs:///output`으로 대몬(daemon) 로그를 복사합니다. 다음 명령을 실행합니다.

  - `--cluster-id`는 클러스터를 지정합니다.
  - `Jar`은 S3DistCp JAR 파일의 위치입니다. command-runner.jar을 사용하여 클러스터에서 명령을 실행하는 방법의 예제는 사용자 지정 JAR 단계를 제출하여 스크립트 또는 명령 실행을 참조하세요.
  - `Args`는 S3DistCp로 전달할 옵션 이름-값 페어의 쉼표로 구분된 목록입니다. 사용 가능한 옵션의 전체 목록은 S3DistCp 옵션 단원을 참조하십시오.

  S3DistCp 복사 단계를 실행 중인 클러스터에 추가하려면 Amazon S3 또는 로컬 파일 시스템에 저장된 JSON 파일(이 예제의 경우 *mySteр.json*)에 다음을 추가합니다. *j-3GYXXXXXX9IOK*를 클러스터 ID로 바꾸고 *amzn-s3-demo-bucket*을 Amazon S3 버킷 이름으로 바꿉니다.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-
demo-bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*[a-
zA-Z,]+"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
```

```
]
```

```
aws emr add-steps --cluster-id j-3GYXXXXXX9IOK --steps file://./myStep.json
```

Example Amazon S3에서 HDFS로 로그 파일 복사

이 예제에서는 실행 중인 클러스터에 단계를 추가하여 Amazon S3 버킷에 저장된 로그 파일을 HDFS로 복사하는 방법을 보여줍니다. 이 예제에서 `--srcPattern` 옵션은 복사된 데이터를 데몬 로그로 제한하는 데 사용됩니다.

`--srcPattern` 옵션을 사용하여 Amazon S3에서 HDFS로 로그 파일을 복사하려면 Amazon S3 또는 로컬 파일 시스템에 저장된 JSON 파일(이 예제의 경우 *myStep.json*)에 다음을 추가합니다. *j-3GYXXXXXX9IOK*를 클러스터 ID로 바꾸고 *amzn-s3-demo-bucket*을 Amazon S3 버킷 이름으로 바꿉니다.

```
[
    {
        "Name":"S3DistCp step",
        "Args":["s3-dist-cp","--s3Endpoint=s3.amazonaws.com","--src=s3://amzn-s3-demo-
bucket/logs/j-3GYXXXXXX9IOJ/node/","--dest=hdfs:///output","--srcPattern=.*daemons.*-
hadoop-.*"],
        "ActionOnFailure":"CONTINUE",
        "Type":"CUSTOM_JAR",
        "Jar":"command-runner.jar"
    }
]
```

실패한 S3DistCp 작업 후 정리

S3DistCp가 지정된 파일의 일부 또는 전체를 복사할 수 없는 경우 명령 또는 클러스터 단계가 실패하고 0 이외의 오류 코드를 반환합니다. 이러한 상황이 발생할 경우 S3DistCp는 부분적으로 복사된 파일을 지우지 않습니다. 수동으로 삭제해야 합니다.

부분적으로 복사된 파일은 S3DistCP 작업의 고유 식별자를 사용하여 하위 디렉터리의 HDFS `tmp` 디렉터리에 저장됩니다. 이 ID는 작업의 표준 출력에서 찾을 수 있습니다.

예를 들어 ID가 `4b1c37bb-91af-4391-aaf8-46a6067085a6`인 S3DistCP 작업의 경우 클러스터의 마스터 노드에 연결하고 다음 명령을 실행하여 작업과 연관된 출력 파일을 볼 수 있습니다.

```
hdfs dfs -ls /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6/output
```

이 명령은 다음과 유사한 파일 목록을 반환합니다.

```
Found 8 items
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/_SUCCESS
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00000
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00001
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:02 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00002
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00003
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00004
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00005
-rw-r--r--   1 hadoop hadoop          0 2018-12-10 06:03 /tmp/4b1c37bb-91af-4391-
aaf8-46a6067085a6/output/part-r-00006
```

그 후에 다음 명령을 실행하여 디렉터리와 모든 내용을 삭제할 수 있습니다.

```
hdfs dfs rm -rf /tmp/4b1c37bb-91af-4391-aaf8-46a6067085a6
```

Amazon EMR 클러스터에서 명령 및 스크립트 실행

이 주제에서는 클러스터에서 명령 또는 스크립트를 단계별로 실행하는 방법을 다룹니다. 명령이나 스크립트를 단계별로 실행하는 방법은 클러스터에 작업을 제출할 수 있는 여러 방법 중 하나이며 다음과 같은 상황에서 유용합니다.

- Amazon EMR 클러스터에 대한 SSH 액세스 권한이 없는 경우
- bash 또는 shell 명령을 실행하여 클러스터 문제를 해결하려는 경우

클러스터를 생성하거나 클러스터가 `WAITING` 상태일 때 스크립트를 실행할 수 있습니다. 단계 처리가 시작되기 전에 스크립트를 실행하려면 부트스트랩 작업을 대신 사용합니다. 부트스트랩 작업에 대한 자세한 내용은 Amazon EMR 관리 안내서에서 부트스트랩 작업을 생성하여 추가 소프트웨어 설치를 참조하세요.

Amazon EMR은 스크립트, 명령 및 기타 클러스터 내 프로그램을 실행하는 데 도움이 되는 다음과 같은 도구를 제공합니다. Amazon EMR 관리 콘솔 또는 AWS CLI를 사용하여 두 도구를 모두 간접 호출할 수 있습니다.

`command-runner.jar`

클러스터의 Amazon EMR AMI에 있습니다. `command-runner.jar`을 사용하여 클러스터에서 명령을 실행할 수 있습니다. 전체 경로를 사용하지 않고 `command-runner.jar`을 지정합니다.

`script-runner.jar`

`s3://<region>.elasticmapreduce/libs/script-runner/script-runner.jar`에서 Amazon S3에서 호스팅됩니다. *<region>*은 Amazon EMR 클러스터가 위치한 리전입니다. `script-runner.jar`을 사용하여 클러스터에서 로컬로 또는 Amazon S3에 저장된 스크립트를 실행할 수 있습니다. 단계를 제출할 때 `script-runner.jar`의 전체 URI를 지정해야 합니다.

사용자 지정 JAR 단계를 제출하여 스크립트 또는 명령을 실행합니다.

다음 AWS CLI 예제에서는 Amazon EMR`script-runner.jar`에서 `command-runner.jar` 및의 몇 가지 일반적인 사용 사례를 보여줍니다.

Example : **command-runner.jar**을 사용하여 클러스터에서 명령 실행

command-runner.jar을 사용할 때 단계의 인수 목록에서 명령, 옵션 및 값을 지정합니다.

다음 AWS CLI 예제에서는를 호출하는 실행 중인 클러스터에 단계를 제출합니다command-runner.jar. Args 목록에 지정된 명령은 Amazon S3에서 *my-script.sh* 스크립트를 hadoop 사용자 홈 디렉터리로 다운로드합니다. 그러면 이 명령은 스크립트의 권한을 수정하고 *my-script.sh*를 실행합니다.

를 사용하는 경우 Args 목록 AWS CLI의 항목은 목록 요소 사이에 공백 없이 쉼표로 구분되어야 합니다. 예를 들어 Args=[example-command, example-option, "example option value"] 대신 Args=[example-command,example-option,"example option value"] 항목을 사용합니다.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Download a script from S3, change its permissions, and
 run it",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[bash,-c,"aws s3 cp s3://
amzn-s3-demo-bucket/my-script.sh /home/hadoop; chmod u+x /home/hadoop/my-script.sh;
 cd /home/hadoop; ./my-script.sh"]
```

Example : **script-runner.jar**을 사용하여 클러스터에서 스크립트 실행

script-runner.jar을 사용할 때 단계의 인수 목록에서 실행하려는 스크립트를 지정합니다.

다음 AWS CLI 예제에서는를 호출하는 실행 중인 클러스터에 단계를 제출합니다script-runner.jar. 이 경우 *my-script.sh* 스크립트는 Amazon S3에 저장됩니다. 클러스터의 프라이머리 노드에 저장되는 로컬 스크립트를 지정할 수도 있습니다.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Run a script from S3 with script-
runner.jar",ActionOnFailure=CONTINUE,Jar=s3://us-west-2.elasticmapreduce/libs/script-
runner/script-runner.jar,Args=[s3://amzn-s3-demo-bucket/my-script.sh]
```

command-runner.jar을 사용하는 기타 방법

command-runner.jar을 사용하여 spark-submit 또는 hadoop-streaming 등의 도구로 클러스터에 작업을 제출할 수 있습니다. command-runner.jar을 사용하여 애플리케이션을 시작할 때

SPARK, STREAMING 또는 PIG와 같은 값을 사용하는 대신 단계 유형으로 CUSTOM_JAR을 지정합니다. 도구 가용성은 클러스터에 설치한 애플리케이션에 따라 달라집니다.

다음 예제 명령에서는 command-runner.jar을 사용하여 spark-submit을 통해 단계를 제출합니다. Args 목록은 spark-submit을 명령으로 지정하고 그 뒤에 Spark 애플리케이션 *my-app.py*의 Amazon S3 URI와 인수 및 값을 포함합니다.

```
aws emr add-steps \
--cluster-id j-2AXXXXXXGAPLF \
--steps Type=CUSTOM_JAR,Name="Run spark-submit using command-
runner.jar",ActionOnFailure=CONTINUE,Jar=command-runner.jar,Args=[spark-submit,S3://
amzn-s3-demo-bucket/my-app.py,ArgName1,ArgValue1,ArgName2,ArgValue2]
```

다음 테이블에는 command-runner.jar을 사용하여 실행할 수 있는 추가 도구가 나와 있습니다.

도구 이름	설명
hadoop-streaming	Hadoop 스트리밍 프로그램을 제출하니다. 콘솔과 일부 SDK에서 이 작업은 스트리밍 단계입니다.
hive-script	Hive 스크립트를 실행합니다. 콘솔과 일부 SDK에서 이 작업은 Hive 단계입니다.
pig-script	Pig 스크립트를 실행합니다. 콘솔과 일부 SDK에서 이 작업은 Pig 단계입니다.
spark-submit	Spark 애플리케이션을 실행합니다. 콘솔에서 이 작업은 Spark 단계입니다.
hadoop-lzo	디렉터리에서 Hadoop LZO 인덱서를 실행합니다.
s3-dist-cp	Amazon S3의 대량 데이터를 HDFS로 분산 복사합니다. 자세한 내용은 S3DistCp(s3-dist-cp) 단원을 참조하십시오.

AWS 용어집

최신 AWS 용어는 AWS 용어집 참조의 AWS 용어집을 참조하세요.