Guía de administración

Amazon EMR

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Amazon EMR: Guía de administración

Copyright © 2025 Amazon Web Services, Inc. and/or its affiliates. All rights reserved.

Las marcas comerciales y la imagen comercial de Amazon no se pueden utilizar en relación con ningún producto o servicio que no sea de Amazon, de ninguna manera que pueda causar confusión entre los clientes y que menosprecie o desacredite a Amazon. Todas las demás marcas registradas que no son propiedad de Amazon son propiedad de sus respectivos propietarios, que pueden o no estar afiliados, conectados o patrocinados por Amazon.

Table of Contents

¿Qué es Amazon EMR? ... 1
Entender cómo crear clústeres de Amazon EMR y trabajar con ellos ... 1
Familiarizarse con los clústeres y los nodos ... 2
Envío de trabajo a un clúster ... 2
Procesamiento de datos ... 3
Descripción del ciclo de vida del clúster ... 4
Beneficios de usar Amazon EMR ... 5
Ahorro de costos ... 6
AWS integración ... 7
Implementación ... 7
Escalabilidad y flexibilidad ... 8
Fiabilidad ... 8
Seguridad ... 9
Monitorización ... 11
Interfaces de administración ... 11
Arquitectura y capas de servicio ... 12
Almacenamiento ... 12
Administración de recursos de clúster ... 13
Marcos de procesamiento de datos ... 14
Aplicaciones y programas ... 15
Antes de configurar Amazon EMR ... 16
Inscríbase en una Cuenta de AWS ... 16
Creación de un usuario con acceso administrativo ... 16
Crear un par de EC2 claves de Amazon para SSH ... 18
Pasos a seguir a continuación ... 18
Tutorial introductorio ... 19
Configuración del clúster de Amazon EMR ... 19
Paso 1: Configurar recursos de datos y lanzar un clúster de Amazon EMR ... 20
Prepare el almacenamiento para Amazon EMR ... 20
Preparar una aplicación con datos de entrada para Amazon EMR ... 20
Lanzar un clúster de Amazon EMR ... 23
Paso 2: Enviar el trabajo a su clúster de Amazon EMR ... 26
Envío de trabajos y visualización de resultados ... 26
Ver los resultados ... 29

Paso 3: Limpieza ... 33
Terminar su clúster ... 33
Eliminar recursos de S3 ... 35
Pasos a seguir a continuación ... 36
Explorar las aplicaciones de macrodatos para Amazon EMR ... 36
Planificar el hardware, las redes y la seguridad de los clústeres ... 36
Administrar clústeres ... 36
Utilizar una interfaz diferente ... 36
Consultar el blog técnico de EMR ... 37
Administración de clústeres de Amazon EMR con la consola ... 38
Capacidades de la consola ... 38
Resumen de las diferencias ... 39
Compatibilidad de clústeres en la consola ... 39
Creación de clústeres ... 39
Visualización y búsqueda de clústeres ... 41
Visualización o edición de los detalles del clúster ... 42
Diferencias al trabajar con configuraciones de seguridad ... 43
Amazon EMR Studio ... 45
Características principales ... 45
IDE de Amazon SageMaker Unified Studio ... 46
Historial de características ... 46
Funcionamiento ... 47
Autenticación e inicio de sesión de los usuarios ... 48
Control de acceso ... 52
Workspaces ... 53
Almacenamiento de cuadernos ... 54
Características, requisitos y límites de EMR Studio ... 54
Consideraciones ... 54
Problemas conocidos ... 57
Limitaciones de características ... 59
Límites de los servicios ... 59
Prácticas recomendadas sobre VPC y subredes para EMR Studio ... 60
Requisitos del clúster ... 60
Configuración de EMR Studio ... 63
Permisos de administrador para crear un EMR Studio ... 63
Configurar un EMR Studio ... 70

Supervise, actualice y elimine los recursos de Amazon EMR Studio 143
Cifrado de las libretas del espacio de trabajo 150
Controlar el tráfico de red de EMR Studio 153
Crear plantillas de clúster 155
Acceso y permisos para los repositorios basados en Git 161
Optimizar los trabajos de Spark 164
Uso de un EMR Studio 165
Conozca los espacios de trabajo de EMR Studio 166
Colaboración en el espacio de trabajo 174
Ejecutar un espacio de trabajo con un rol de tiempo de ejecución 178
Ejecutar los cuadernos de espacios de trabajo de Amazon EMR Studio mediante programación 183
Examinar los datos con SQL Explorer para EMR Studio 184
Asociar computación a un espacio de trabajo 185
Vincular repositorios de Git 193
Integración de Athena 197
CodeWhisperer integración 198
Depurar aplicaciones y trabajos 200
Instalar kernels y bibliotecas 205
Comandos magic 206
Usar cuadernos multilingües con kernels de Spark 215
EMR Notebooks 217
Cuadernos en la consola 218
Acerca de la transición 218
¿Qué necesita hacer? 219
Ventajas del espacio de trabajo 219
Permisos necesarios 220
Consideraciones 221
Requisitos del clúster 221
Diferencias en capacidades por versión de clúster 223
Límites para cuadernos asociados de forma simultánea 224
Versiones de cuaderno de Jupyter y Python 224
Consideraciones en torno a la seguridad 225
Crear un bloc de notas 225
Uso de los Cuadernos de EMR 229
Descripción de los estados del cuaderno 229

Uso del editor de cuadernos ... 231
Cambio de clústeres ... 232
Eliminación de cuadernos y archivos de cuadernos ... 233
Uso compartido de archivos de cuadernos ... 234
Ejemplos de comandos programáticos para Cuadernos de EMR ... 235
Descripción general ... 235
Permisos ... 236
Limitaciones ... 237
Ejemplos ... 238
Ejemplos de comandos de la CLI de Cuadernos ... 238
Ejemplo de script del SDK de Boto3 ... 244
Script de ejemplo de Ruby ... 246
Suplantación de usuarios para Spark ... 249
Configuración de la suplantación de usuarios de Spark ... 249
Uso del widget de supervisión de trabajos de Spark ... 250
Seguridad ... 250
Instalación y uso de kernels y bibliotecas en EMR Studio ... 251
... 252
Instalación de kernels y bibliotecas de Python en un nodo principal del clúster ... 252
Consideraciones y limitaciones de las bibliotecas adaptadas al cuaderno ... 255
Uso de bibliotecas adaptadas al cuaderno ... 256
Asociación de repositorios basados en Git con Cuadernos de EMR ... 257
Requisitos y consideraciones previos ... 258
Agregar un repositorio basado en Git a Amazon EMR ... 261
Actualizar o eliminar un repositorio basado en Git de un espacio de trabajo de EMR Studio ... 262
Vincular o desvincular un repositorio basado en Git en EMR Studio ... 263
Crear un nuevo cuaderno con un repositorio de Git asociado en EMR Studio ... 265
Uso de repositorios de Git en un Cuaderno de EMR Studio ... 266
Planificación, configuración y lanzamiento de clústeres de Amazon EMR ... 267
Lanzamiento de un clúster de Amazon EMR ... 267
Configuración de la ubicación del clúster de Amazon EMR y el almacenamiento de datos ... 268
Elija una AWS región para su clúster de Amazon EMR ... 269
Trabajo con almacenamiento y sistemas de archivos ... 270
Preparación de los datos de entrada para su procesamiento con Amazon EMR ... 275
Cómo configurar una ubicación para la salida del clúster de Amazon EMR ... 295

Planificación y configuración de nodos principales en su clúster de Amazon EMR 302
Características que admiten la alta disponibilidad en un clúster de Amazon EMR y cómo funcionan con aplicaciones de código abierto 303
Lanzar un clúster de Amazon EMR con varios nodos principales 313
Integración de Amazon EMR con grupos de ubicación EC2 319
Consideraciones y prácticas recomendadas al crear un clúster con varios nodos principales 327
EMR se agrupa en AWS Outposts 330
Requisitos previos 330
Limitaciones 330
Consideraciones sobre la conectividad de red 331
Crear un clúster de Amazon EMR en AWS Outposts 332
Clústeres de EMR en Zonas Locales AWS 333
Tipos de instancias admitidas 333
Creación de un clúster de Amazon EMR en zonas locales 334
Configuración de Docker para su uso con clústeres de Amazon EMR 335
Registros de Docker 336
Configuración de registros de Docker 337
Configuración de YARN para acceder a Amazon ECR en EMR 6.0.0 y versiones anteriores 338
Control de la finalización de clústeres de Amazon EMR 340
Configuración de un clúster de Amazon EMR para que continúe o finalice después de la ejecución de pasos 341
Uso de una política de terminación automática 344
Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental 350
Sustitución de nodos en mal estado con Amazon EMR 357
Configuración predeterminada de protección de reemplazo y finalización de nodos 358
Configuración de la sustitución de nodos en mal estado al lanzar un clúster 359
Configuración de la sustitución de nodos en mal estado en un clúster en ejecución 360
Trabajando con AMIs 362
Descripción general 362
Uso de la AMI predeterminada 363
Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR 541
Cambio de versión de AL 555

Personalización del volumen raíz de EBS ... 556
Configure las aplicaciones al lanzar su clúster de Amazon EMR ... 559
Creación de acciones de arranque ... 560
Configuración del hardware y las redes de los clústeres de Amazon EMR ... 565
Comprender los tipos de nodos ... 566
Configurar los tipos de EC2 instancias de Amazon para utilizarlos con Amazon EMR ... 570
Configuración del registro y la depuración de un clúster de Amazon EMR ... 1552
Archivos de registro predeterminados ... 1553
Archivar archivos de registro en Amazon S3 ... 1554
Localización de registros ... 1560
Etiquetado y clasificación de los recursos del clúster de Amazon EMR ... 1561
Restricciones que se aplican al etiquetado de recursos en Amazon EMR ... 1563
Etiquetado de recursos de Amazon EMR para facturación ... 1564
Cómo añadir etiquetas a un clúster de Amazon EMR ... 1564
Visualización de las etiquetas de un clúster de Amazon EMR ... 1566
Eliminación de las etiquetas de un clúster de Amazon EMR ... 1567
Integración de controladores y aplicaciones de terceros en Amazon EMR ... 1569
Utilizar herramientas de inteligencia empresarial con Amazon EMR ... 1569
Seguridad ... 1570
Seguridad de la red y la infraestructura ... 1570
Actualizaciones de la AMI de Amazon Linux ... 1571
AWS Identity and Access Management con Amazon EMR ... 1572
Clústeres de un solo inquilino y de varios inquilinos ... 1573
Protección de los datos ... 1574
Control de acceso a los datos ... 1574
Configuraciones de seguridad ... 1575
Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI ... 1575
Cómo especificar una configuración de seguridad ... 1607
Protección de los datos ... 1608
Cifrar datos en reposo y en tránsito con Amazon EMR ... 1609
IAM con Amazon EMR ... 1647
Público ... 1648
Autenticación con identidades ... 1648
Administración de acceso mediante políticas ... 1652
Cómo funciona Amazon EMR con IAM ... 1655
Roles en tiempo de ejecución para los pasos de Amazon EMR ... 1662

Configuración de roles de servicio para Amazon EMR ... 1672
Ejemplos de políticas basadas en identidades ... 1739
S3 Access Grants con Amazon EMR ... 1784
Descripción general ... 1784
Funcionamiento ... 1785
Consideraciones ... 1786
Lance un clúster ... 1787
`Lake Formation` ... 1788
`fallbackToIAM` ... 1789
Autenticación en nodos de clúster ... 1789
Utilice un par de EC2 claves para las credenciales SSH de Amazon EMR ... 1790
Uso de la autenticación de Kerberos ... 1790
Uso de la autenticación LDAP ... 1825
Integración de Amazon EMR con Identity Center ... 1837
Descripción general ... 1837
Características ... 1837
Introducción ... 1838
Consideraciones ... 1847
Integración de Amazon EMR con Lake Formation ... 1848
Cómo funciona Amazon EMR con Lake Formation ... 1849
Requisitos previos ... 1849
Habilitación de Lake Formation con Amazon EMR ... 1850
Hudi y Lake Formation ... 1854
Iceberg y Lake Formation ... 1857
Delta Lake y Lake Formation ... 1858
Consideraciones ... 1860
Uso de vistas del catálogo de datos de Glue en Amazon EMR ... 1863
Uso de un rol de tiempo de ejecución de tareas de Lake Formation con acceso completo a las tablas ... 1870
Integración de Amazon EMR con Apache Ranger ... 1877
Información general de Ranger ... 1878
Consideraciones sobre el uso de Amazon EMR con Apache Ranger ... 1880
Configuración de Amazon EMR para Apache Ranger ... 1882
Complementos Apache Ranger para escenarios de integración de Amazon EMR ... 1900
Solución de problemas con Apache Ranger ... 1919
Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR ... 1923

Uso de grupos de seguridad administrados por Amazon EMR 1926
Trabajo con grupos de seguridad adicionales para un clúster de Amazon EMR 1937
Especificación de los grupos de seguridad 1938
Grupos de seguridad para Cuadernos de Amazon EMR 1941
Bloqueo de acceso público 1943
Validación de conformidad 1948
Resiliencia 1949
Seguridad de la infraestructura 1950
Conexión a Amazon EMR mediante un punto de conexión de VPC de tipo interfaz 1951
Administración de clústeres de Amazon EMR 1956
Conectar a un clúster de Amazon EMR 1956
Antes de conectarse 1957
Conexión al nodo principal del clúster de Amazon EMR mediante SSH 1959
Envío del trabajo a un clúster de Amazon EMR 1985
Agregar pasos con la consola 1986
Adición de pasos con la CLI 1989
Ejecutar varios pasos 1991
Visualización de los pasos después de enviar el trabajo a un clúster de Amazon EMR 1992
Cancelación de pasos al enviar trabajo a un clúster de Amazon EMR 1993
Visualización y monitorización de un clúster de Amazon EMR a medida que realiza su trabajo 1995
Visualización del estado y los detalles del clúster de Amazon EMR 1995
Depuración de pasos mejorada con Amazon EMR 2001
Visualización del historial de aplicaciones de Amazon EMR 2003
Visualización de los archivos de registro de Amazon EMR 2013
Ver instancias de clúster en Amazon EC2 2017
CloudWatch eventos y métricas de Amazon EMR 2018
Visualización de métricas de aplicaciones de clúster mediante Ganglia con Amazon EMR 2111
CloudTrail registros 2111
Mejores prácticas de observabilidad de EMR 2116
Utilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes 2117
Consideraciones 2118
Escalado administrado 2119
Escalado automático con una política personalizada 2163
Cambiar el tamaño de un clúster en ejecución 2176

Tiempos de espera de aprovisionamiento 2184
Opciones de reducción vertical de clústeres para clústeres de Amazon EMR 2189
Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera. 2193
Terminación desde la consola 2194
Terminación desde la CLI 2194
Terminación desde la API 2196
Clonar un clúster 2196
Automatización de los clústeres de Amazon EMR recurrentes con AWS Data Pipeline 2197
Tutoriales de Amazon EMR 2198
Amazon EMR activado EC2 : supervisión mejorada con el CloudWatch uso de métricas y registros personalizados 2198
Descripción general 2198
Requisitos previos y antecedentes 2198
..... 2199
Supervise las aplicaciones de Apache Spark en Amazon EMR con Amazon CloudWatch 2209
Supervise el estado de la solicitud de Amazon EMR con la integración CloudWatch 2209
Solución de problemas de clústeres de Amazon EMR 2211
Herramientas de solución de problemas 2211
Ver detalles del clúster 2212
Ver detalles de errores 2212
Ejecutar scripts y configurar procesos 2213
Ver archivos de registro de 2213
Supervisar el rendimiento del clúster 2214
Ver y reiniciar procesos 2214
Ver procesos en ejecución 2215
Detener y reiniciar procesos 2216
Errores comunes 2219
Códigos de error 2220
Errores de recursos durante las operaciones del clústeres de Amazon EMR 2238
Errores de entrada y salida de clúster durante la operaciones de Amazon EMR 2253
Errores de permisos durante las operaciones del clúster de Amazon EMR 2256
Errores de clúster de Hive 2257
Errores de la VPC durante las operaciones del clúster de Amazon EMR 2259
Errores del clúster de Amazon EMR de streaming 2263
Amazon EMR: errores de clúster JAR personalizados 2265
Errores de Amazon EMR AWS GovCloud (EE. UU. - Oeste) 2265

Buscar un clúster que falta ... 2266
resolución de problemas de clústeres con errores ... 2266
Paso 1: recopile datos sobre el problema con el clúster de Amazon EMR ... 2267
Paso 2: comprobar el entorno ... 2268
Paso 3: comprobar el último cambio de estado ... 2269
Paso 4: examine los archivos de registro de Amazon EMR ... 2269
Paso 5: compruebe el clúster de Amazon EMR de forma pormenorizada ... 2271
Solucionar problemas de un clúster lento ... 2272
Paso 1: recopile datos sobre el problema con el clúster de Amazon EMR ... 2273
Paso 2: compruebe el entorno del clúster de Amazon EMR ... 2273
Paso 3: examine los archivos de registro del clúster de Amazon EMR ... 2275
Paso 4: compruebe el estado de la instancia y el clúster de Amazon EMR ... 2277
Paso 5: comprobar si hay grupos suspendidos ... 2278
Paso 6: revise los ajustes de configuración del clúster de Amazon EMR ... 2279
Paso 7: examine los datos de entrada del clúster de Amazon EMR ... 2282
Solución de problemas comunes al utilizar Amazon EMR AWS con Lake Formation ... 2282
No se permite el acceso al lago de datos ... 2282
Vencimiento de la sesión ... 2283
No hay permisos para el usuario en la tabla solicitada ... 2283
Consulta de datos entre cuentas compartidos con Lake Formation ... 2283
Inserción, creación y alteración de tablas ... 2285
Escritura de aplicaciones que lanzan y administran clústeres de Amazon EMR ... 2286
End-to-end Ejemplo de código fuente Java de Amazon EMR ... 2286
Conceptos comunes para las llamadas a la API de Amazon EMR ... 2290
Puntos de conexión para Amazon EMR ... 2291
Especificar parámetros de clúster en Amazon EMR ... 2291
Zonas de disponibilidad en Amazon EMR ... 2292
Cómo utilizar archivos y bibliotecas adicionales en clústeres de Amazon EMR ... 2292
Se usa SDKs para llamar a Amazon EMR APIs ... 2293
Uso de AWS SDK para Java para crear un clúster de Amazon EMR ... 2293
Administrar las Service Quotas de Amazon EMR ... 2296
¿Qué son las Service Quotas de Amazon EMR? ... 2296
Cómo administrar Service Quotas para Amazon EMR ... 2297
Cuándo configurar los eventos de EMR en CloudWatch ... 2298
AWS Glosario ... 2301
Historial de documentos ... 2302

..... mmcccciii

¿Qué es Amazon EMR?

Amazon EMR, que anteriormente se denominaba Amazon Elastic MapReduce, es una plataforma de clústeres gestionada que simplifica la ejecución de marcos de big data, como Apache Hadoop y Apache Spark, AWS para procesar y analizar grandes cantidades de datos. Mediante el uso de estos marcos de trabajo y proyectos de código abierto relacionados, puede procesar datos para fines de análisis y cargas de trabajo de inteligencia empresarial. Además, Amazon EMR le permite transformar y trasladar grandes cantidades de datos hacia y desde otros almacenes de datos y bases de datos de AWS , tales como Amazon Simple Storage Service (Amazon S3) y Amazon DynamoDB.

Si está utilizando Amazon EMR por primera vez, le recomendamos que comience leyendo lo siguiente, además de esta sección:

- Amazon EMR: esta página de servicio ofrece los puntos destacados de Amazon EMR, los detalles del producto y la información de precios.
- Tutorial: Introducción a Amazon EMR: este tutorial le ayudará a empezar a utilizar Amazon EMR rápidamente.

En esta sección

- Entender cómo crear clústeres de Amazon EMR y trabajar con ellos
- Beneficios de usar Amazon EMR
- Arquitectura y capas de servicio en Amazon EMR

Entender cómo crear clústeres de Amazon EMR y trabajar con ellos

En este tema, se ofrece información general de los clústeres de Amazon EMR, lo que incluye cómo enviar trabajos a un clúster, la forma en que se procesan los datos y los distintos estados por los que pasa el clúster durante el procesamiento.

En este tema

- Familiarizarse con los clústeres y los nodos
- Envío de trabajo a un clúster

- Procesamiento de datos
- Descripción del ciclo de vida del clúster

Familiarizarse con los clústeres y los nodos

El componente central de Amazon EMR es el clúster. Un clúster es un conjunto de instancias de Amazon Elastic Compute Cloud (Amazon EC2). Cada instancia del clúster se denomina nodo. Cada nodo tiene un rol dentro del clúster, conocido como el tipo de nodo. Amazon EMR también instala distintos componentes de software en cada tipo de nodo, lo que proporciona a cada nodo un rol en una aplicación distribuida, como Apache Hadoop.

Los tipos de nodos en Amazon EMR son los siguientes:

- Nodo principal: un nodo que administra el clúster mediante la ejecución de componentes de software para coordinar la distribución de datos y las tareas entre los demás nodos para su procesamiento. El nodo principal hacer un seguimiento del estado de las tareas y supervisa el estado del clúster. Cada clúster tiene un nodo principal y se puede crear un clúster de un solo nodo con solo el nodo principal.
- Nodo secundario: un nodo con componentes de software que ejecutan tareas y almacenan datos en el Hadoop Distributed File System (HDFS) del clúster. Los clústeres de varios nodos tienen al menos un nodo secundario.
- Nodo de tareas: un nodo con componentes de software que solo ejecuta tareas y no almacena datos en HDFS. Los nodos de tareas son opcionales.

Envío de trabajo a un clúster

Cuando se ejecuta un clúster en Amazon EMR, dispone de varias opciones sobre cómo especificar el trabajo que hay que llevar a cabo.

- Proporcionar toda la definición del trabajo que hay que realizar en funciones que debe especificar como pasos al crear un clúster. Esto se realiza normalmente para clústeres que procesan una cantidad definida de datos y, a continuación, terminan cuando se completa el procesamiento.
- Cree un clúster de ejecución prolongada y utilice la consola de Amazon EMR, la API de Amazon EMR o AWS CLI los pasos para enviar, que pueden contener uno o más trabajos. Para obtener más información, consulte Envío del trabajo a un clúster de Amazon EMR.

- Cree un clúster, conéctese al nodo principal y a los demás nodos según sea necesario mediante SSH y utilice las interfaces que proporcionan las aplicaciones instaladas para llevar a cabo tareas y enviar consultas, ya sea de forma interactiva o con scripts. Para obtener más información, consulte la Guía de publicación de Amazon EMR.

Procesamiento de datos

Al lanzar el clúster, puede elegir los marcos de trabajo y las aplicaciones que desea instalar para sus necesidades de procesamiento de datos. Para procesar datos en el clúster de Amazon EMR, puede enviar los trabajos o las consultas directamente a las aplicaciones instaladas, o puede ejecutar pasos en el clúster.

Envío de trabajos directamente a las aplicaciones

Puede enviar trabajos e interactuar directamente con el software que está instalado en el clúster de Amazon EMR. Para ello, normalmente se conecta al nodo principal a través de una conexión segura y accede a las interfaces y herramientas que están disponibles para que el software se ejecute directamente en el clúster. Para obtener más información, consulte Conectar a un clúster de Amazon EMR.

Ejecución de pasos para procesar datos

Puede enviar uno o varios pasos ordenados a un clúster de Amazon EMR. Cada paso es una unidad de trabajo que contiene instrucciones para manipular los datos para su procesamiento por el software instalado en el clúster.

A continuación se muestra un proceso de ejemplo que utiliza cuatro pasos:

1. Enviar un conjunto de datos de entrada para procesamiento.
2. Procesar la salida del primer paso mediante un programa de Pig.
3. Procesar un segundo conjunto de datos de entrada mediante un programa de Hive.
4. Escribir un conjunto de datos de salida.

Por lo general, cuando se procesan datos en Amazon EMR, la entrada son datos almacenados como archivos en el sistema de archivos subyacente elegido, como Amazon S3 o HDFS. Estos datos se transfieren de un paso al siguiente en la secuencia de procesamiento. El último paso escribe los datos de salida en una ubicación especificada, como un bucket de Amazon S3.

Los pasos se ejecutan en la siguiente secuencia:

1. Se envía una solicitud para empezar los pasos de procesamiento.
2. El estado de todos los pasos se establece en PENDING (Pendiente).
3. Cuando se inicia el primer paso de la secuencia, su estado cambia a RUNNING (En ejecución). Los demás pasos permanecen en el estado PENDING (Pendiente).
4. Una vez que finaliza el primer paso, su estado cambia a COMPLETED (Completado).
5. El siguiente paso de la secuencia se inicia y su estado cambia a RUNNING (En ejecución). Una vez finalizado, su estado cambia a COMPLETED (Completado).
6. Este patrón se repite para cada paso hasta que se completen todos y el procesamiento finaliza.

El siguiente diagrama representa la secuencia de pasos y cambios de estado de los pasos a medida que se procesan.

Si un paso falla durante el procesamiento, su estado cambia a ERROR. Puede determinar lo que ocurre para cada paso. De forma predeterminada, los pasos restantes de la secuencia se establecen en CANCELADO y no se ejecutan si falla un paso anterior. También puede elegir omitir el error y permitir que los pasos restantes continúen o terminar el clúster inmediatamente.

El siguiente diagrama representa la secuencia de pasos y el cambio de estado predeterminado cuando un paso produce un error durante el procesamiento.

Descripción del ciclo de vida del clúster

Un clúster de Amazon EMR correcto sigue este proceso:

1. Amazon EMR primero aprovisiona EC2 instancias en el clúster para cada instancia según sus especificaciones. Para obtener más información, consulte Configuración del hardware y las redes de los clústeres de Amazon EMR. Para todas las instancias, Amazon EMR utiliza la AMI predeterminada de Amazon EMR o la AMI de Amazon Linux personalizada que se especifique. Para obtener más información, consulte Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR. Durante esta fase, el estado del clúster es `STARTING`.
2. Amazon EMR ejecuta las acciones de arranque que especifique en cada instancia. Puede utilizar acciones de arranque para instalar aplicaciones personalizadas y realizar las personalizaciones

que necesite. Para obtener más información, consulte Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR. Durante esta fase, el estado del clúster es `BOOTSTRAPPING`.

3. Amazon EMR instala las aplicaciones nativas que especifique al crear el clúster, tales como Hive, Hadoop, Spark, etc.
4. Cuando las acciones de arranque se han completado correctamente y las aplicaciones nativas se han instalado correctamente, el estado del clúster es `RUNNING`. En este punto, puede conectarse a las instancias del clúster y el clúster ejecutará por orden todos los pasos que haya especificado al crear el clúster. Puede enviar pasos adicionales, que se ejecutarán después de los pasos anteriores. Para obtener más información, consulte Envío del trabajo a un clúster de Amazon EMR.
5. Una vez que los pasos se han ejecutado correctamente, el clúster pasa al estado `WAITING`. Si un clúster está configurado para que se termine automáticamente después de que se haya completado el último paso, pasa a un estado `TERMINATING` y luego al estado `TERMINATED`. Si el clúster está configurado para esperar, debe apagarlo manualmente cuando ya no lo necesite. Después de terminar manualmente el clúster, pasa al estado `TERMINATING` y, a continuación, al estado `TERMINATED`.

Un error durante el ciclo de vida del clúster hace que Amazon EMR termine el clúster y todas sus instancias, a menos que habilite la protección de terminación. Si se termina un clúster debido a un error, todos los datos almacenados en el clúster se eliminan y el estado del clúster se establece en `TERMINATED_WITH_ERRORS`. Si ha habilitado la protección de terminación, puede recuperar los datos del clúster y después eliminar la protección de terminación y terminar el clúster. Para obtener más información, consulte Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental.

El siguiente diagrama representa el ciclo de vida de un clúster y cómo cada etapa del ciclo de vida se asocia a un determinado estado del clúster.

Beneficios de usar Amazon EMR

El uso de Amazon EMR conlleva muchos beneficios. Estos incluyen la flexibilidad que ofrece AWS y los ahorros de costos disponibles en comparación con la creación de sus propios recursos locales. En esta sección se ofrece información general sobre estos beneficios y enlaces a información adicional que le ayudan a seguir explorando.

Temas

- Ahorro de costos
- AWS integración
- Implementación
- Escalabilidad y flexibilidad
- Fiabilidad
- Seguridad
- Monitorización
- Interfaces de administración

Ahorro de costos

Los precios de Amazon EMR dependen del tipo de instancia y del número de EC2 instancias de Amazon que implemente y de la región en la que lance el clúster. Los precios bajo demanda ofrecen tarifas reducidas, pero puede reducir aún más el costo mediante la adquisición de instancias reservadas o instancias de spot. Las instancias de spot pueden ofrecer ahorros significativos. En algunos casos, de solo una décima parte de los precios bajo demanda.

> **Note**
>
> Si utiliza Amazon S3, Amazon Kinesis o DynamoDB con el clúster de EMR, hay cargos adicionales por estos servicios que se facturan por separado de su uso de Amazon EMR.

> **Note**
>
> Al configurar un clúster de Amazon EMR en una subred privada, se recomienda que también configure puntos de conexión de VPC para Amazon S3. Si su clúster de EMR se encuentra en una subred privada sin puntos de conexión de VPC para Amazon S3, incurrirá en cargos adicionales de puerta de enlace de NAT asociados al tráfico de S3, ya que el tráfico entre su clúster de EMR y S3 no permanecerá dentro de su VPC.

Para obtener más información sobre las opciones de precios y los detalles, consulte Precios de Amazon EMR.

AWS integración

Amazon EMR se integra con otros AWS servicios para proporcionar capacidades y funcionalidades relacionadas con las redes, el almacenamiento, la seguridad, etc., para su clúster. La siguiente lista proporciona diversos ejemplos de esta integración:

- Amazon EC2 para las instancias que componen los nodos del clúster
- Amazon Virtual Private Cloud (Amazon VPC) para configurar la red virtual en la que lanzar sus instancias
- Amazon S3 para almacenar los datos de entrada y de salida
- Amazon supervisará CloudWatch el rendimiento de los clústeres y configurará las alarmas
- AWS Identity and Access Management (IAM) para configurar los permisos
- AWS CloudTrail para auditar las solicitudes realizadas al servicio
- AWS Data Pipeline para programar e iniciar sus clústeres
- AWS Lake Formation para descubrir, catalogar y proteger los datos en un lago de datos de Amazon S3

Implementación

El clúster de EMR se compone de EC2 instancias que realizan el trabajo que usted envía a su clúster. Al lanzar el clúster, Amazon EMR configura las instancias con las aplicaciones que elija, como Apache Hadoop o Spark. Elija el tamaño y el tipo de instancia que mejor se adapte a las necesidades de procesamiento de su clúster: procesamiento por lotes, consultas de baja latencia, streaming de datos o almacenamiento de datos de gran tamaño. Para obtener más información sobre los tipos de instancias disponibles para Amazon EMR, consulte Configuración del hardware y las redes de los clústeres de Amazon EMR.

Amazon EMR ofrece distintas formas de configurar el software en su clúster. Por ejemplo, puede instalar una versión de Amazon EMR con un conjunto seleccionado de aplicaciones que puede incluir marcos de trabajo versátiles como, por ejemplo, Hadoop y aplicaciones como Hive, Pig o Spark. También puede instalar una de las diversas distribuciones de MapR. Amazon EMR utiliza Amazon Linux, por lo que también puede instalar software en su clúster de forma manual o mediante el administrador de paquetes yum o desde el origen. Para obtener más información, consulte Configure las aplicaciones al lanzar su clúster de Amazon EMR.

Escalabilidad y flexibilidad

Amazon EMR proporciona flexibilidad para aumentar o reducir el escalado del clúster a medida que cambien sus necesidades informáticas. Puede cambiar el tamaño del clúster para añadir instancias durante los picos de cargas de trabajo y eliminar instancias para controlar los costos cuando desaparezcan los picos de cargas de trabajo. Para obtener más información, consulte Cambio manual del tamaño de un clúster de Amazon EMR en ejecución.

Amazon EMR también ofrece la opción de ejecutar varios grupos de instancias, para que puede utilizar instancias bajo demanda en un grupo para garantizar la potencia de procesamiento junto con las instancias de spot en otro grupo para completar los trabajos con mayor rapidez y para reducir costos. También puede combinar diversos tipos de instancias para aprovechar mejor los precios de un tipo de instancia de spot sobre otros. Para obtener más información, consulte ¿Cuándo se deben utilizar las instancias de spot?.

Además, Amazon EMR proporciona la flexibilidad necesaria para utilizar diversos sistemas de archivos para sus datos de entrada, de salida e intermedios. Por ejemplo, podría elegir el Sistema de archivos distribuido de Hadoop (HDFS), que se ejecuta en los nodos principal y secundarios del clúster para el procesamiento de datos que no es necesario almacenar más allá del ciclo de vida del clúster. Podría elegir el sistema de archivos de EMR (EMRFS) para utilizar Amazon S3 como capa de datos para aplicaciones que se ejecutan en su clúster para que pueda separar la computación y el almacenamiento, así como para conservar los datos cuando finaliza el ciclo de vida de su clúster. EMRFS ofrece el beneficio añadido de permitirle aumentar o reducir el escalado de sus necesidades de computación y almacenamiento de manera independiente. Para escalar sus necesidades de computación, puede cambiar el tamaño de su clúster y puede escalar sus necesidades de almacenamiento mediante el uso de Amazon S3. Para obtener más información, consulte Trabajo con almacenamiento y sistemas de archivos con Amazon EMR.

Fiabilidad

Amazon EMR supervisa los nodos del clúster y termina y sustituye automáticamente una instancia en caso de que se produzca algún error.

Amazon EMR proporciona opciones de configuración que controlan cómo se termina el clúster: de forma automática o manual. Si configura el clúster para terminarlo de forma automática, se termina después de que se completen todos los pasos. Esto se conoce como clúster transitorio. Sin embargo, puede configurar el clúster para que se siga ejecutando una vez completado el procesamiento, para que pueda elegir terminarlo manualmente cuando ya no lo necesite. O bien,

puede crear un clúster, interactuar con las aplicaciones instaladas directamente y, a continuación, terminar el clúster manualmente cuando deje de necesitarlo. Los clústeres de estos ejemplos se denominan clústeres de ejecución prolongada.

Además, puede configurar la protección de terminación para evitar que las instancias del clúster se terminen por errores o problemas durante el procesamiento. Cuando se habilita la protección de terminación, puede recuperar los datos desde instancias antes de la terminación. La configuración predeterminada de estas opciones varía en función de si lanza el clúster a través de la consola, la CLI o la API. Para obtener más información, consulte Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental.

Seguridad

Amazon EMR aprovecha otros AWS servicios, como IAM y Amazon VPC, y funciones como los pares de claves de EC2 Amazon, para ayudarlo a proteger sus clústeres y datos.

IAM

Amazon EMR se integra con IAM para administrar los permisos. Puede definir permisos mediante las políticas de IAM, que se asocian a usuarios o grupos de IAM. Los permisos que defina en la política determinan las acciones que dichos usuarios o miembros del grupo pueden realizar y los recursos a los que pueden acceder. Para obtener más información, consulte Cómo funciona Amazon EMR con IAM.

Además, Amazon EMR utiliza las funciones de IAM para el propio servicio de Amazon EMR y el perfil de EC2 instancia para las instancias. Estas funciones otorgan permisos para que el servicio y las instancias accedan a otros AWS servicios en su nombre. Hay un rol predeterminado para el servicio Amazon EMR y un rol predeterminado para el perfil de la EC2 instancia. Los roles predeterminados utilizan políticas AWS administradas, que se crean automáticamente la primera vez que se lanza un clúster de EMR desde la consola y se eligen los permisos predeterminados. También puede crear los roles de IAM predeterminados desde la AWS CLI. Si desea administrar los permisos en lugar de hacerlo AWS, puede elegir roles personalizados para el perfil de servicio e instancia. Para obtener más información, consulte Configuración de los roles de servicio de IAM de los permisos de Amazon EMR para los servicios y recursos de AWS.

Grupos de seguridad

Amazon EMR utiliza grupos de seguridad para controlar el tráfico entrante y saliente a sus instancias. EC2 Al lanzar el clúster, Amazon EMR utiliza un grupo de seguridad para la instancia principal y un

grupo de seguridad que las core/task instances. Amazon EMR configures the security group rules to ensure communication among the instances in the cluster. Optionally, you can configure additional security groups and assign them to your primary and core/task instancias deben compartir para establecer reglas más avanzadas. Para obtener más información, consulte Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

Cifrado

Amazon EMR admite cifrado del cliente y del servidor opcional de Amazon S3 con EMRFS para ayudar a proteger los datos que se almacenan en Amazon S3. Con el cifrado del lado del servidor, Amazon S3 cifra sus datos después de cargarlos.

Con el cifrado del lado cliente, el proceso de cifrado y descifrado se produce en el cliente EMRFS en su clúster de EMR. Usted administra la clave raíz para el cifrado del lado del cliente mediante AWS Key Management Service (AWS KMS) o su propio sistema de administración de claves.

Para obtener más información, consulte Especificación del cifrado de Amazon S3 con propiedades de EMRFS.

Amazon VPC

Amazon EMR admite el lanzamiento de clústeres en una nube privada virtual (VPC) en Amazon VPC. Una VPC es una red virtual aislada AWS que permite controlar aspectos avanzados de la configuración y el acceso a la red. Para obtener más información, consulte Configuración de redes en una VPC para Amazon EMR.

AWS CloudTrail

Amazon EMR se integra CloudTrail para registrar la información sobre las solicitudes realizadas por su cuenta o en su AWS nombre. Con esta información, puede realizar un seguimiento de quién accede en cada momento a su clúster y la dirección IP desde la que se ha realizado la solicitud. Para obtener más información, consulte Registro de llamadas a la API AWS EMR mediante AWS CloudTrail.

Pares de EC2 claves de Amazon

Para supervisar e interactuar con el clúster, puede formar una conexión segura entre el equipo remoto y el nodo principal. Puede utilizar el protocolo de red Secure Shell (SSH) para esta conexión o Kerberos para la autenticación. Si utilizas SSH, necesitarás un par de EC2 claves de Amazon.

Para obtener más información, consulte Utilice un par de EC2 claves para las credenciales SSH de Amazon EMR.

Monitorización

Puede utilizar las interfaces de administración de Amazon EMR y los archivos de registro para solucionar problemas de clúster como, por ejemplo, averías o errores. Amazon EMR ofrece la posibilidad de archivar los archivos de registro en Amazon S3 para que pueda almacenar registros y cuestiones de solución de problemas incluso después de que el clúster termine. Amazon EMR también proporciona una herramienta de depuración opcional en la consola de Amazon EMR para examinar los archivos de registro basada en pasos, trabajos y tareas. Para obtener más información, consulte Configuración del registro y la depuración de un clúster de Amazon EMR.

Amazon EMR se integra CloudWatch para realizar un seguimiento de las métricas de rendimiento del clúster y de los trabajos dentro del clúster. Puede configurar alarmas basadas en diversas métricas como, por ejemplo, si el clúster está inactivo o el porcentaje de almacenamiento utilizado. Para obtener más información, consulte Supervisión de las métricas de Amazon EMR con CloudWatch.

Interfaces de administración

Existen varias formas en las que puede interactuar con Amazon EMR:

- Consola: una interfaz gráfica de usuario que puede utilizar para lanzar y administrar clústeres. Con ella, puede rellenar formularios web para especificar los detalles de los clústeres que lanzar, ver los detalles de clústeres existentes, depurar y terminar clústeres. El uso de la consola es la manera más sencilla de empezar a utilizar Amazon EMR; no se requieren conocimientos de programación. La consola está disponible en línea en https://console.aws.amazon.com/elasticmapreduce/casa.
- AWS Command Line Interface (AWS CLI): una aplicación cliente que ejecuta en su máquina local para conectarse a Amazon EMR y crear y administrar clústeres. AWS CLI Contiene un conjunto de comandos rico en funciones específicos de Amazon EMR. Con ella, puede escribir scripts que automatizan el proceso de lanzamiento y administración de clústeres. Si prefiere trabajar desde una línea de comandos, la mejor opción AWS CLI es utilizar la. Para obtener más información, consulte Amazon EMR en la Referencia de los comandos de la AWS CLI .
- Kit de desarrollo de software (SDK): SDKs proporciona funciones que utilizan Amazon EMR para crear y gestionar clústeres. Con ellos, puede escribir aplicaciones que automatizan el proceso de creación y administración de clústeres. Utilizar el SDK es la mejor opción para ampliar o personalizar la funcionalidad de Amazon EMR. Amazon EMR está disponible actualmente en las siguientes versionesSDKs: Go, Java, .NET (C# y VB.NET), Node.js, PHP, Python y Ruby.

Para obtener más información al respecto SDKs, consulte Herramientas para AWS y códigos y bibliotecas de ejemplo de Amazon EMR.

- API de servicios web: un interfaz de bajo nivel que puede utilizar para llamar al servicio web directamente, utilizando JSON. El uso de la API es la mejor opción para crear un SDK personalizado que llame a Amazon EMR. Para obtener más información, consulte la Referencia de las API de Amazon EMR.

Arquitectura y capas de servicio en Amazon EMR

La arquitectura de servicio de Amazon EMR se compone de varias capas, cada una de las cuales proporciona determinadas capacidades y funcionalidad al clúster. Esta sección proporciona información general sobre las capas y los componentes de cada una de ellas.

En este tema

- Almacenamiento
- Administración de recursos de clúster
- Marcos de procesamiento de datos
- Aplicaciones y programas

Almacenamiento

La capa de almacenamiento incluye los diferentes sistemas de archivos que se utilizan con el clúster. Existen varios tipos distintos de opciones de almacenamiento como se indica a continuación.

Sistema de archivos distribuido de Hadoop (HDFS)

Hadoop Distributed File System (HDFS) es un sistema de archivos distribuido y escalable para Hadoop. HDFS distribuye los datos que almacena en instancias en el clúster, almacenando varias copias en datos en distintas instancias para garantizar que no se pierdan datos si una instancia individual falla. HDFS es un almacenamiento efímero que se reclama cuando se termina un clúster. El HDFS es útil para almacenar en caché los resultados intermedios durante el MapReduce procesamiento o para cargas de trabajo que tienen una cantidad significativa de E/S aleatorias.

Para obtener más información, consulte Comportamiento y opciones de almacenamiento de instancias en Amazon EMR en esta guía o vaya a la Guía del usuario de HDFS en el sitio web de Apache Hadoop.

Sistema de archivos de EMR (EMRFS)

Con el sistema de archivos de EMR (EMRFS), Amazon EMR amplía Hadoop para agregar la posibilidad de tener acceso directamente a los datos almacenados en Amazon S3 como si fueran un sistema de archivos similar a HDFS. Puede usar HDFS o Amazon S3 como sistema de archivos en su clúster. A menudo, Amazon S3 se usa para almacenar los datos de entrada y de salida y los resultados intermedios se almacenan en HDFS.

Sistema de archivos local

El sistema de archivos local se refiere a un disco conectado a nivel local. Al crear un clúster de Hadoop, cada nodo se crea a partir de una EC2 instancia de Amazon que incluye un bloque preconfigurado de almacenamiento en disco preconectado denominado almacén de instancias. Los datos de los volúmenes del almacén de instancias solo se conservan durante el ciclo de vida de su EC2 instancia de Amazon.

Administración de recursos de clúster

La capa de administración de recursos es responsable de la administración de los recursos de clúster y de la programación de trabajos para procesamiento de datos.

De forma predeterminada, Amazon EMR utiliza YARN (Yet Another Resource Negotiator), que es un componente introducido en Apache Hadoop 2.0 para administrar de forma centralizada recursos de clúster para varios marcos de procesamiento de datos. Sin embargo, hay otros marcos y aplicaciones que se ofrecen en Amazon EMR que no utilizan YARN como administrador de recursos. Amazon EMR también dispone de un agente en cada nodo que administra los componentes de YARN, mantiene el clúster en buen estado y se comunica con el servicio de Amazon EMR.

Dado que las instancias de spot se utilizan a menudo para ejecutar nodos de tarea, Amazon EMR tiene una funcionalidad predeterminada para programar trabajos de YARN, de modo que los trabajos en ejecución no presenten errores cuando los nodos de tarea que se ejecutan en las instancias de spot se terminen. Para ello, Amazon EMR permite que los procesos maestros de la aplicación se ejecuten únicamente en los nodos principales. El proceso maestro de la aplicación controla los trabajos en ejecución y debe mantenerse activo durante toda la vida del trabajo.

La versión 5.19.0 y posteriores de Amazon EMR utilizan la característica integrada de etiquetas de nodo YARN para lograrlo. (Las versiones anteriores utilizaban una revisión de código). Las propiedades en las clasificaciones de configuración `yarn-site` y `capacity-scheduler` se ajustan de forma predeterminada para que capacity-scheduler y fair-scheduler de YARN utilicen las

etiquetas de nodo. Amazon EMR etiqueta automáticamente los nodos principales con la etiqueta `CORE` y establece las propiedades para que los maestros de la aplicación se programen únicamente en los nodos con la etiqueta CORE. La modificación manual de las propiedades relacionadas en las clasificaciones de configuración yarn-site y capacity-scheduler o directamente en los archivos XML asociados podría interrumpir esta característica o modificar esta funcionalidad.

Marcos de procesamiento de datos

La capa de marco de trabajo de procesamiento de datos es el motor que se utiliza para procesar y analizar datos. Existen muchos marcos de trabajo disponibles que se ejecutan en YARN o que tienen su propia administración de recursos. Los distintos marcos están disponibles para los diferentes tipos de necesidades de procesamiento tales como lotes, interactivo, en memoria, streaming, etc. El marco de trabajo que elija depende de su caso de uso. Esto afecta a los lenguajes y a los interfaces disponibles desde la capa de aplicación, que es la capa que se utiliza para interactuar con los datos que desea procesar. Los principales marcos de procesamiento disponibles para Amazon EMR son MapReduce Hadoop y Spark.

Hadoop MapReduce

Hadoop MapReduce es un modelo de programación de código abierto para la computación distribuida. Simplifica el proceso de escritura de aplicaciones distribuidas en paralelo mediante el tratamiento de toda la lógica, mientras proporciona las funciones Map y Reduce. La función Map asigna datos a conjuntos de pares clave-valor denominados resultados intermedios. La función Reduce combina los resultados intermedios, aplica algoritmos adicionales y genera la salida final. Hay varios marcos disponibles MapReduce, como Hive, que genera automáticamente los programas Map y Reduce.

Para obtener más información, consulte Cómo asignar y reducir operaciones que se están llevando a cabo en el sitio web de la wiki de Apache Hadoop.

Apache Spark

Spark es un marco de trabajo de clúster y un modelo de programación para el procesamiento de cargas de trabajo de big data. Al igual que Hadoop MapReduce, Spark es un sistema de procesamiento distribuido de código abierto, pero utiliza gráficos acíclicos dirigidos para los planes de ejecución y almacenamiento en caché en memoria para los conjuntos de datos. Cuando se ejecuta Spark en Amazon EMR, es posible utilizar EMRFS para tener acceso directamente a los datos en Amazon S3. Spark admite diversos módulos de consulta interactivos como, por ejemplo, SparkSQL.

Para obtener más información, consulte Apache Spark en clústeres de Amazon EMR en la Guía de publicación de Amazon EMR.

Aplicaciones y programas

Amazon EMR es compatible con muchas aplicaciones, tales como Hive, Pig y la biblioteca Spark Streaming, para ofrecer diversas capacidades, como el uso de lenguajes de nivel superior para crear cargas de trabajo de procesamiento, el uso de algoritmos de machine learning, el desarrollo de aplicaciones de procesamiento de flujos y la creación de almacenamientos de datos. Además, Amazon EMR también admite proyectos de código abierto que tienen su propia funcionalidad de administración de clústeres en lugar de utilizar YARN.

Puede utilizar diversas bibliotecas y lenguajes para interactuar con las aplicaciones que se ejecutan en Amazon EMR. Por ejemplo, puedes usar Java, Hive o Pig con MapReduce Spark Streaming, Spark SQL y GraphX con Spark. MLlib

Para obtener más información, consulte la Guía de publicación de Amazon EMR.

Antes de configurar Amazon EMR

Complete las tareas preliminares detalladas en esta sección antes de lanzar un clúster de Amazon EMR por primera vez. Estas incluyen configurar su AWS cuenta si la necesita y tomar medidas para configurar una comunicación segura.

Inscríbase en una Cuenta de AWS

Si no tiene uno Cuenta de AWS, complete los siguientes pasos para crearlo.

Para suscribirse a una Cuenta de AWS

1. Abrir https://portal.aws.amazon.com/billing/registro.
2. Siga las instrucciones que se le indiquen.

   Parte del procedimiento de registro consiste en recibir una llamada telefónica o mensaje de texto e indicar un código de verificación en el teclado del teléfono.

   Cuando te registras en un Cuenta de AWS, Usuario raíz de la cuenta de AWSse crea un. El usuario raíz tendrá acceso a todos los Servicios de AWS y recursos de esa cuenta. Como práctica recomendada de seguridad, asigne acceso administrativo a un usuario y utilice únicamente el usuario raíz para realizar tareas que requieren acceso de usuario raíz.

AWS te envía un correo electrónico de confirmación una vez finalizado el proceso de registro. En cualquier momento, puede ver la actividad de su cuenta actual y administrarla accediendo a https://aws.amazon.com/y seleccionando Mi cuenta.

Creación de un usuario con acceso administrativo

Después de crear un usuario administrativo Cuenta de AWS, asegúrelo Usuario raíz de la cuenta de AWS AWS IAM Identity Center, habilite y cree un usuario administrativo para no usar el usuario root en las tareas diarias.

Proteja su Usuario raíz de la cuenta de AWS

1. Inicie sesión AWS Management Consolecomo propietario de la cuenta seleccionando el usuario root e introduciendo su dirección de Cuenta de AWS correo electrónico. En la siguiente página, escriba su contraseña.

Para obtener ayuda para iniciar sesión con el usuario raíz, consulte Iniciar sesión como usuario raíz en la Guía del usuario de AWS Sign-In .

2. Active la autenticación multifactor (MFA) para el usuario raíz.

   Para obtener instrucciones, consulte Habilitar un dispositivo MFA virtual para el usuario Cuenta de AWS raíz (consola) en la Guía del usuario de IAM.

Creación de un usuario con acceso administrativo

1. Activar IAM Identity Center.

   Consulte las instrucciones en Activar AWS IAM Identity Center en la Guía del usuario de AWS IAM Identity Center .

2. En IAM Identity Center, conceda acceso administrativo a un usuario.

   Para ver un tutorial sobre su uso Directorio de IAM Identity Center como fuente de identidad, consulte Configurar el acceso de los usuarios con la configuración predeterminada Directorio de IAM Identity Center en la Guía del AWS IAM Identity Center usuario.

Inicio de sesión como usuario con acceso de administrador

- Para iniciar sesión con el usuario de IAM Identity Center, use la URL de inicio de sesión que se envió a la dirección de correo electrónico cuando creó el usuario de IAM Identity Center.

  Para obtener ayuda para iniciar sesión con un usuario del Centro de identidades de IAM, consulte Iniciar sesión en el portal de AWS acceso en la Guía del AWS Sign-In usuario.

Concesión de acceso a usuarios adicionales

1. En IAM Identity Center, cree un conjunto de permisos que siga la práctica recomendada de aplicar permisos de privilegios mínimos.

   Para conocer las instrucciones, consulte Create a permission set en la Guía del usuario de AWS IAM Identity Center .

2. Asigne usuarios a un grupo y, a continuación, asigne el acceso de inicio de sesión único al grupo.

Para conocer las instrucciones, consulte Add groups en la Guía del usuario de AWS IAM Identity Center .

Crear un par de EC2 claves de Amazon para SSH

> **Note**
>
> Con la versión 5.10.0 o las versiones posteriores de Amazon EMR, puede configurar Kerberos para autenticar a los usuarios y las conexiones SSH con el clúster. Para obtener más información, consulte Uso de Kerberos para la autenticación con Amazon EMR.

Para autenticarse y conectarse a los nodos de un clúster a través de un canal seguro mediante el protocolo Secure Shell (SSH), cree un par de claves de Amazon Elastic Compute Cloud (Amazon EC2) antes de lanzar el clúster. También puede crear un clúster sin un par de claves. Esto se hace con clústeres transitorios que se inician, ejecutan pasos, y luego se terminan de forma automática.

Si…	Entonces...
Ya tienes un par de EC2 claves de Amazon que quieres usar o no necesitas autenticarte en tu clúster.	Omita este paso.
Debe crear un par de claves.	Consulta Cómo crear tu key pair con Amazon EC2.

Pasos a seguir a continuación

- Para obtener información sobre cómo crear un clúster de muestra, consulte Tutorial: Introducción a Amazon EMR.
- Para obtener más información sobre cómo configurar un clúster personalizado y controlar el acceso al mismo, consulte Planificación, configuración y lanzamiento de clústeres de Amazon EMR y Seguridad en Amazon EMR.

Tutorial: Introducción a Amazon EMR

Revise un flujo de trabajo para configurar rápidamente un clúster de Amazon EMR y ejecutar una aplicación de Spark.

Configuración del clúster de Amazon EMR

Amazon EMR, puede configurar un clúster para procesar y analizar datos con marcos de macrodatos en solo unos minutos. En este tutorial, se muestra cómo lanzar un clúster de muestra con Spark y cómo ejecutar un PySpark script sencillo almacenado en un bucket de Amazon S3. Cubre tareas esenciales de Amazon EMR en tres categorías principales de flujo de trabajo: planificación y configuración, administración y limpieza.

Encontrará enlaces a temas más detallados a medida que avance en el tutorial e ideas para pasos adicionales en la sección Pasos a seguir a continuación. Si tiene alguna pregunta o no sabe cómo continuar, contacte con el equipo de Amazon EMR en nuestro foro de discusión.

Requisitos previos

- Antes de lanzar un clúster de Amazon EMR, asegúrese de completar las tareas en Antes de configurar Amazon EMR.

Costo

- El clúster de ejemplo que cree se ejecuta en un entorno real. El clúster genera unos gastos mínimos. Para evitar cargos adicionales, asegúrese de completar las tareas de limpieza del último paso de este tutorial. Los cargos se acumulan en virtud la tarifa por segundo según los precios de Amazon EMR. Los cargos también varían según la región. Para obtener más información, consulte Precios de Amazon EMR.
- Es posible que se acumulen cargos mínimos por los archivos pequeños que almacene en Amazon S3. Es posible que no se apliquen algunos o todos los cargos de Amazon S3 si se encuentra dentro de los límites de uso de la capa AWS gratuita. Para obtener más información, consulte Precios de Amazon S3 y Nivel gratuito de AWS.

Paso 1: Configurar recursos de datos y lanzar un clúster de Amazon EMR

Prepare el almacenamiento para Amazon EMR

Cuando utiliza Amazon EMR, puede elegir entre una variedad de sistemas de archivos para almacenar los datos de entrada, los datos de salida y los archivos de registro. En este tutorial, utilizará EMRFS para almacenar datos en un bucket de S3. EMRFS es una implementación del sistema de archivos de Hadoop que le permite leer y escribir archivos normales en Amazon S3. Para obtener más información, consulte Trabajo con almacenamiento y sistemas de archivos con Amazon EMR.

Para crear un bucket para este tutorial, consulte Crear un bucket de S3 en la Guía del usuario de la consola de Amazon Simple Storage Service. Cree el bucket en la misma AWS región en la que planea lanzar su clúster de Amazon EMR. Por ejemplo, Oeste de EE. UU. (Oregón) us-west-2.

Los buckets y las carpetas que utilice con Amazon EMR tienen las siguientes limitaciones:

- Los nombres pueden contener letras minúsculas, números, puntos (.) y guiones (-).
- Los nombres no pueden terminar en números.
- El nombre de un bucket debe ser único en todas las cuentas de AWS .
- La carpeta de salida debe estar vacía.

Preparar una aplicación con datos de entrada para Amazon EMR

La forma más común de preparar una aplicación para Amazon EMR consiste en cargar la aplicación y sus datos de entrada en Amazon S3. A continuación, cuando envíe el trabajo a su clúster, especifique las ubicaciones de Amazon S3 para el script y los datos.

En este paso, debe cargar un PySpark script de muestra en su bucket de Amazon S3. Le proporcionamos un PySpark script para que lo utilice. El script procesa los datos de inspección del establecimiento de alimentos y devuelve un archivo de resultados en su bucket de S3. En el archivo de resultados se enumeran los diez establecimientos con más infracciones de tipo “rojo”.

También debe cargar datos de entrada de muestra en Amazon S3 para que el PySpark script los procese. Los datos de entrada son una versión modificada de los resultados de las inspecciones del Departamento de Salud en el condado de King (Washington) entre 2006 y 2020. Para obtener más

información, consulte Datos abiertos del condado de King: datos de inspección de establecimientos alimentarios. No descargues los datos de los restaurantes para este tutorial directamente del sitio web del condado de King, ya que es un archivo muy grande. A continuación incluimos una descarga que contiene un archivo con menos registros para ayudarte a completar el tutorial. A continuación se incluyen ejemplos de filas del conjunto de datos.

```
name,inspection_result,inspection_closed_business,violation_type,violation_points
100 LB CLAM,Unsatisfactory,FALSE,BLUE,5
100 PERCENT NUTRICION,Unsatisfactory,FALSE,BLUE,5
7-ELEVEN #2361-39423A,Complete,FALSE,,0
```

Para preparar el PySpark script de ejemplo para EMR

1. Copie el siguiente código de muestra en un nuevo archivo en el editor que prefiera.

```
import argparse

from pyspark.sql import SparkSession

def calculate_red_violations(data_source, output_uri):
    """
    Processes sample food establishment inspection data and queries the data to
 find the top 10 establishments
    with the most Red violations from 2006 to 2020.

    :param data_source: The URI of your food establishment data CSV, such as 's3://
amzn-s3-demo-bucket/food-establishment-data.csv'.
    :param output_uri: The URI where output is written, such as 's3://amzn-s3-demo-
bucket/restaurant_violation_results'.
    """
    with SparkSession.builder.appName("Calculate Red Health
 Violations").getOrCreate() as spark:
        # Load the restaurant violation CSV data
        if data_source is not None:
            restaurants_df = spark.read.option("header", "true").csv(data_source)

        # Create an in-memory DataFrame to query
        restaurants_df.createOrReplaceTempView("restaurant_violations")

        # Create a DataFrame of the top 10 restaurants with the most Red violations
        top_red_violation_restaurants = spark.sql("""SELECT name, count(*) AS
 total_red_violations
```

```
          FROM restaurant_violations
          WHERE violation_type = 'RED'
          GROUP BY name
          ORDER BY total_red_violations DESC LIMIT 10""")

        # Write the results to the specified output URI
        top_red_violation_restaurants.write.option("header",
 "true").mode("overwrite").csv(output_uri)

if __name__ == "__main__":
    parser = argparse.ArgumentParser()
    parser.add_argument(
        '--data_source', help="The URI for you CSV restaurant data, like an S3
 bucket location.")
    parser.add_argument(
        '--output_uri', help="The URI where output is saved, like an S3 bucket
 location.")
    args = parser.parse_args()

    calculate_red_violations(args.data_source, args.output_uri)
```

2. Guarde el archivo como `health_violations.py`.
3. Suba `health_violations.py` a Amazon S3 en el bucket que creó para este tutorial. Para obtener instrucciones, consulte Cargar un objeto en un bucket en la Guía de introducción a Amazon Simple Storage Service.

Para preparar los datos de entrada de muestra para EMR

1. Descargue el archivo zip, food_establishment_data.zip.
2. Descomprima y guarde `food_establishment_data.zip` como `food_establishment_data.csv` en su máquina.
3. Suba el archivo CSV en el bucket de S3 que creó para este tutorial. Para obtener instrucciones, consulte Cargar un objeto en un bucket en la Guía de introducción a Amazon Simple Storage Service.

Para obtener más información sobre cómo configurar los datos para EMR, consulte Preparación de los datos de entrada para su procesamiento con Amazon EMR.

Lanzar un clúster de Amazon EMR

Tras preparar la ubicación de almacenamiento y la aplicación, puede lanzar un clúster de muestra de Amazon EMR. En este paso, lanza un clúster de Apache Spark con la versión de Amazon EMR más reciente.

Console

Para lanzar un clúster con Spark instalado con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En la página Crear clúster, anote los valores predeterminados para la versión, el tipo de instancia, el número de instancias y los permisos. Estos campos se rellenan automáticamente con valores que funcionan para clústeres de uso general.
4. En el campo Nombre del clúster, introduzca un nombre de clúster único que le ayude a identificar el clúster, por ejemplo. `My first cluster` El nombre del clúster no puede contener los caracteres <, >, $, | o `(acento grave).
5. En Aplicaciones, elije la opción Spark para instalar Spark en el clúster.

   > **Note**
   >
   > Elija las aplicaciones que desee incluir en su clúster de Amazon EMR antes de lanzarlo. No puede agregar ni eliminar aplicaciones de un clúster después del lanzamiento.

6. En Registros de clúster, seleccione la casilla Publicar los registros específicos del clúster en Amazon S3. Sustituya el valor de Ubicación de Amazon S3 por el bucket de Amazon S3 que creó, seguido por **`/logs`**. Por ejemplo, **`s3://amzn-s3-demo-bucket/logs`**. Al agregar **`/logs`**, se crea una nueva carpeta llamada “logs” en el bucket, donde Amazon EMR puede copiar los archivos de registro del clúster.
7. En Configuración de seguridad y permisos, elige tu EC2 key pair. En la misma sección, selecciona el menú desplegable Función de servicio para Amazon EMR y selecciona EMR_. DefaultRole A continuación, seleccione el menú desplegable de la función de IAM, por ejemplo, el perfil, y elija EMR_ _. EC2 DefaultRole

8. Seleccione Crear clúster para lanzar el clúster y abrir la página de detalles del clúster.
9. Consulte el estado del clúster junto a su nombre. El estado cambia de Iniciando a En ejecución a Esperando a medida que Amazon EMR aprovisiona el clúster. Es posible que tenga que elegir el icono de actualización situado a la derecha o actualizar el navegador para ver las actualizaciones de estado.

Cuando el estado cambia a Esperando, el clúster está activo, en ejecución y listo para aceptar trabajo. Para obtener más información sobre cómo leer el resumen del clúster, consulte Visualización del estado y los detalles del clúster de Amazon EMR. Para obtener más información acerca del estado del clúster, consulte Descripción del ciclo de vida del clúster.

CLI

Para lanzar un clúster con Spark instalado con el AWS CLI

1. Cree los roles predeterminados de IAM que, a continuación, podrá utilizar para crear el clúster mediante el siguiente comando.

```
aws emr create-default-roles
```

   Para obtener más información sobre `create-default-roles`, consulte la Referencia de los comandos de la AWS CLI.

2. Cree un clúster de Spark con el siguiente comando. Introduzca un nombre para el clúster con la `--name` opción y especifique el nombre de su EC2 key pair con la `--ec2-attributes` opción.

```
aws emr create-cluster \
--name "<My First EMR Cluster>" \
--release-label <emr-5.36.2> \
--applications Name=Spark \
--ec2-attributes KeyName=<myEMRKeyPairName> \
--instance-type m5.xlarge \
--instance-count 3 \
--use-default-roles
```

   Anote los demás valores obligatorios para `--instance-type`, `--instance-count`, y `--use-default-roles`. Estos valores se han elegido para clústeres de uso general. Para

obtener más información sobre `create-cluster`, consulte la Referencia de los comandos de la AWS CLI.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Debería ver un resultado como el siguiente. El resultado muestra el `ClusterId` y el `ClusterArn` del nuevo clúster. Anote su `ClusterId`. Se utiliza `ClusterId` para comprobar el estado del clúster y enviar trabajos.

```
{
    "ClusterId": "myClusterId",
    "ClusterArn": "myClusterArn"
}
```

3. Compruebe el estado del clúster con el siguiente comando.

```
aws emr describe-cluster --cluster-id <myClusterId>
```

Debería ver un resultado como el siguiente con el objeto `Status` del nuevo clúster.

```
{
    "Cluster": {
        "Id": "myClusterId",
        "Name": "My First EMR Cluster",
        "Status": {
            "State": "STARTING",
            "StateChangeReason": {
                "Message": "Configuring cluster software"
            }
        }
     }
}
```

El valor `State` cambia de `STARTING` a `RUNNING` a `WAITING` cuando Amazon EMR aprovisiona el clúster.

El estado del clúster cambia a **`WAITING`** cuando un clúster está activo, en ejecución y listo para aceptar trabajo. Para obtener más información acerca del estado del clúster, consulte Descripción del ciclo de vida del clúster.

Paso 2: Enviar el trabajo a su clúster de Amazon EMR

Envío de trabajos y visualización de resultados

Después de lanzar un clúster, puede enviar el trabajo al clúster en ejecución para procesar y analizar los datos. El trabajo se envía a un clúster de Amazon EMR como un paso. Un paso es una unidad de trabajo compuesta por una o más acciones. Por ejemplo, puede enviar un paso para calcular valores o para transferir y procesar datos. Puede enviar pasos cuando al crear un clúster o bien a un clúster en ejecución. En este tutorial, envía `health_violations.py` como un paso a su clúster en ejecución. Para obtener más información sobre los pasos, consulte Envío del trabajo a un clúster de Amazon EMR.

Console

Para enviar una aplicación de Spark como paso con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, seleccione el clúster al que desea enviar el trabajo. El estado del clúster debe estar en Esperando.
3. Seleccione la pestaña Pasos y, a continuación Agregar paso.
4. Configure el paso de acuerdo con las directrices siguientes:
   - En Tipo, seleccione Aplicación de Spark. Deberá ver campos adicionales para el modo implementación, la ubicación de la aplicación y las opciones de spark-submit.
   - En Nombre, escriba un nuevo nombre. Si tiene muchos pasos en un clúster, nombrar cada paso le ayudará a hacer un seguimiento de ellos.
   - Para el modo implementación, deje el valor predeterminado Modo de clúster. Para obtener más información sobre los modos de implementación de Spark, consulte Información general sobre el modo de clúster en la documentación de Apache Spark.

- En Ubicación de la aplicación, introduzca la ubicación del `health_violations.py` script en Amazon S3, por ejemplo*`s3://amzn-s3-demo-bucket/health_violations.py`*.
- Deje vacío el campo Opciones de Spark-submit. Para obtener más información sobre las opciones de `spark-submit`, consulte Launching applications with spark-submit.
- En el campo Argumentos, ingrese los siguientes argumentos y valores:

```
--data_source s3://amzn-s3-demo-bucket/food_establishment_data.csv
--output_uri s3://amzn-s3-demo-bucket/myOutputFolder
```

  *`s3://amzn-s3-demo-bucket/food_establishment_data.csv`*Sustitúyalo por el URI del bucket de S3 de los datos de entrada en los que ha preparadoPreparar una aplicación con datos de entrada para Amazon EMR.

  *`amzn-s3-demo-bucket`*Reemplácelo por el nombre del depósito que creó para este tutorial y *`myOutputFolder`* sustitúyalo por un nombre para la carpeta de salida del clúster.

- En Acción si se produce un error en el paso, acepte la opción predeterminada Continuar. De esta forma, si se produce un error en el paso, el clúster seguirá ejecutándose.

5. Seleccione Agregar para enviar el paso. El paso debería aparecer en la consola con el estado Pendiente.
6. Supervise el estado del paso. Debería cambiar de Pendiente a En ejecución a Completado. Para actualizar el estado en la consola, elija el icono de actualización situado a la derecha de Filtro. El script tarda aproximadamente un minuto en ejecutarse. Cuando el paso se haya completado correctamente, el estado cambiará a Completado.

CLI

Para enviar una solicitud de Spark como un paso junto con el AWS CLI

1. Asegúrese de tener el `ClusterId` del clúster que lanzó en Lanzar un clúster de Amazon EMR. Puede recuperar el ID del clúster con los siguientes comandos.

```
aws emr list-clusters --cluster-states WAITING
```

2. Envíe `health_violations.py` como un paso con el comando `add-steps` y el `ClusterId`.

- Puedes especificar un nombre para tu paso sustituyéndolo *"My Spark Application"*. En la `Args` matriz, *s3://amzn-s3-demo-bucket/health_violations.py* sustitúyalo por la ubicación de `health_violations.py` la aplicación.
- *s3://amzn-s3-demo-bucket/food_establishment_data.csv*Sustitúyala por la ubicación S3 de tu `food_establishment_data.csv` conjunto de datos.
- *s3://amzn-s3-demo-bucket/MyOutputFolder*Sustitúyala por la ruta S3 del depósito designado y un nombre para la carpeta de salida del clúster.
- `ActionOnFailure=CONTINUE` significa que el clúster seguirá ejecutándose si se produce un error en el paso.

```
aws emr add-steps \
--cluster-id <myClusterId> \
--steps Type=Spark,Name="<My Spark
 Application>",ActionOnFailure=CONTINUE,Args=[<s3://amzn-s3-demo-
bucket/health_violations.py>,--data_source,<s3://amzn-s3-demo-bucket/
food_establishment_data.csv>,--output_uri,<s3://amzn-s3-demo-bucket/
MyOutputFolder>]
```

Para obtener más información acerca envío de pasos mediante la CLI, consulte la Referencia de los comandos de la AWS CLI.

Después de enviar el paso, debería ver un resultado como el siguiente con una lista de `StepIds`. Como ha enviado un paso, solo verá un ID en la lista. Copie el ID de su paso. Utilice el identificador del paso para comprobar el estado.

```
{
    "StepIds": [
        "s-1XXXXXXXXXXA"
    ]
}
```

3. Consulte el estado del paso con el comando `describe-step`.

```
aws emr describe-step --cluster-id <myClusterId> --step-id <s-1XXXXXXXXXXA>
```

Debería ver un resultado como el siguiente, con información sobre el paso.

```
{
    "Step": {
        "Id": "s-1XXXXXXXXXXA",
        "Name": "My Spark Application",
        "Config": {
            "Jar": "command-runner.jar",
            "Properties": {},
            "Args": [
                "spark-submit",
                "s3://amzn-s3-demo-bucket/health_violations.py",
                "--data_source",
                "s3://amzn-s3-demo-bucket/food_establishment_data.csv",
                "--output_uri",
                "s3://amzn-s3-demo-bucket/myOutputFolder"
            ]
        },
        "ActionOnFailure": "CONTINUE",
        "Status": {
            "State": "COMPLETED"
        }
    }
}
```

A medida que el paso se ejecuta, el `State` cambia de `PENDING` a `RUNNING` a `COMPLETED`. El paso tarda aproximadamente un minuto en ejecutarse, por lo que es posible que tenga que comprobar el estado varias veces.

Sabrá que el paso se ha completado correctamente cuando el `State` cambie a **`COMPLETED`**.

Para obtener más información sobre el ciclo de vida del paso, consulte Ejecución de pasos para procesar datos.

Ver los resultados

Cuando un paso se ejecute correctamente, podrá ver los resultados de salida en la carpeta de salida de Amazon S3.

Para ver los resultados de **`health_violations.py`**

1. Abra la consola de Amazon S3 en https://console.aws.amazon.com/s3/.

2. Elija el nombre del bucket y, a continuación, la carpeta de salida que especificó al enviar el paso. Por ejemplo, *amzn-s3-demo-bucket* y luego*myOutputFolder*.
3. Compruebe que los siguientes elementos aparecen en la carpeta de salida:
   - Un objeto de tamaño pequeño llamado `_SUCCESS`.
   - Un archivo CSV que comienza con el prefijo `part-` que contiene los resultados.
4. Elija el objeto con los resultados y, a continuación, seleccione Descargar para guardar los resultados en el sistema de archivos local.
5. Abra los resultados con el editor que prefiera. En el archivo de salida, se muestra una lista de los diez establecimientos de alimentación con más infracciones rojas. En el archivo de salida, también se muestra el número total de infracciones rojas de cada establecimiento.

   A continuación se muestra un resultado de `health_violations.py`.

```
name, total_red_violations
SUBWAY, 322
T-MOBILE PARK, 315
WHOLE FOODS MARKET, 299
PCC COMMUNITY MARKETS, 251
TACO TIME, 240
MCDONALD'S, 177
THAI GINGER, 153
SAFEWAY INC #1508, 143
TAQUERIA EL RINCONSITO, 134
HIMITSU TERIYAKI, 128
```

Para obtener más información sobre el resultado del clúster de Amazon EMR, consulte Cómo configurar una ubicación para la salida del clúster de Amazon EMR.

(Opcional) Conectarse al clúster de Amazon EMR en ejecución

Cuando utiliza Amazon EMR, es posible que desee conectarse a un clúster en ejecución para leer los archivos de registro, depurar el clúster o utilizar herramientas de la CLI, como el intérprete de comandos de Spark. Amazon EMR le permite conectarse a un clúster mediante el protocolo Secure Shell (SSH). En esta sección, se explica cómo configurar SSH, conectarse al clúster y ver los archivos de registro de Spark. Para obtener más información acerca de la conexión al clúster, consulte Autenticación en nodos de clúster de Amazon EMR.

Autorizar las conexiones de SSH a su clúster

Antes de conectarse al clúster, debe modificar sus grupos de seguridad para autorizar las conexiones SSH entrantes. Los grupos EC2 de seguridad de Amazon actúan como firewalls virtuales para controlar el tráfico entrante y saliente del clúster. Cuando creó el clúster para este tutorial, Amazon EMR creó los siguientes grupos de seguridad en su nombre:

ElasticMapReduce-maestro

El grupo de seguridad administrado por Amazon EMR, asociado al nodo principal. En un clúster de Amazon EMR, el nodo principal es una EC2 instancia de Amazon que administra el clúster.

ElasticMapReduce-esclavo

El grupo de seguridad, asociado a los nodos secundarios y de tareas.

Console

Para permitir el acceso SSH a los orígenes de confianza del grupo de seguridad principal con la consola

Para editar los grupos de seguridad, debe tener permiso para administrar los grupos de seguridad de la VPC en la que se encuentra el clúster. Para obtener más información, consulte Cambiar los permisos de un usuario y el ejemplo de política que permite administrar grupos de EC2 seguridad en la Guía del usuario de IAM.

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija el clúster que desee actualizar. Se abrirá la página de detalles del clúster. La pestaña Propiedades de esta página debe estar preseleccionada.
3. En Redes, en la pestaña Propiedades, seleccione la flecha situada junto a los grupos de EC2 seguridad (firewall) para ampliar esta sección. En Nodo principal, seleccione el enlace del grupo de seguridad. Cuando haya completado los siguientes pasos, si lo desea, puede volver a este paso, elija Nodos principales y de tareas y repita los pasos siguientes para permitir que el cliente SSH acceda a dichos nodos.
4. Esto abre la EC2 consola. Elija la pestaña Reglas de entrada y, a continuación, elija Editar reglas de entrada.

5. Compruebe si hay una regla de entrada que permita el acceso público con la siguiente configuración. Si existe, seleccione Eliminar para eliminarla.

   - Tipo

     SSH
   - Puerto

     22
   - Origen

     0.0.0.0/0 personalizado

   > **Warning**
   >
   > Antes de diciembre de 2020, el grupo de seguridad ElasticMapReduce -master tenía una regla preconfigurada para permitir el tráfico entrante en el puerto 22 desde todas las fuentes. Esta regla se creó para simplificar las conexiones SSH iniciales al nodo maestro. Le recomendamos encarecidamente que elimine esta regla de entrada y que restrinja el tráfico a los orígenes de confianza.

6. Desplácese a la parte inferior de la lista de reglas y seleccione Agregar regla.
7. En Type (Tipo), seleccione SSH. Al seleccionar SSH, se ingresa automáticamente TCP en Protocolo y 22 en Rango de puertos.
8. Como origen, seleccione Mi IP para agregar automáticamente su dirección IP como dirección de origen. También puede agregar un rango de direcciones IP de clientes de confianza personalizadas o crear reglas adicionales para otros clientes. Muchos entornos de red asignan direcciones IP de forma dinámica, por lo que es posible que en el futuro necesite actualizar las direcciones IP de los clientes de confianza.
9. Seleccione Save.
10. Si lo desea, elija Nodos principales y de tareas en la lista y repita los pasos anteriores para permitir que el cliente SSH acceda dichos nodos.

Conéctese a su clúster mediante el AWS CLI

Independientemente del sistema operativo, puede crear una conexión SSH a su clúster mediante la AWS CLI.

Para conectarse a su clúster y ver los archivos de registro mediante el AWS CLI

1. Use el comando siguiente para abrir una conexión SSH a su clúster. *<mykeypair.key>*Sustitúyalo por la ruta completa y el nombre de archivo del archivo de key pair. Por ejemplo, `C:\Users\<username>\.ssh\mykeypair.pem`.

```
aws emr ssh --cluster-id <j-2AL4XXXXXX5T9> --key-pair-file <~/mykeypair.key>
```

2. Navegue hasta `/mnt/var/log/spark` para acceder a los registros de Spark en el nodo maestro del clúster. A continuación, consulte los archivos de esa ubicación. Para obtener una lista de los archivos de registro adicionales del nodo maestro, consulte Ver archivos de registro en el nodo principal.

```
cd /mnt/var/log/spark
ls
```

Use Amazon SageMaker AI Unified Studio para administrar su clúster de Amazon EMR

Amazon EMR on también EC2 es un tipo de procesamiento compatible con Amazon SageMaker AI Unified Studio. Consulte Administración de Amazon EMR EC2 para saber cómo usar y administrar EMR en los EC2 recursos de Unified Studio. Amazon SageMaker AI

Paso 3: eliminar los recursos de Amazon EMR

Terminar su clúster

Ahora que ha enviado el trabajo a su clúster y ha visto los resultados de su PySpark solicitud, puede cancelar el clúster. Al terminar un clúster, se detienen todos los cargos de Amazon EMR y las instancias de Amazon asociados al clúster. EC2

Al terminar un clúster, Amazon EMR conserva los metadatos del clúster durante dos meses sin costo alguno. Los metadatos archivados le ayudan a clonar el clúster para un nuevo trabajo o a revisitar la configuración del clúster como referencia. Los metadatos no incluyen los datos que el clúster escribe en S3 ni los datos almacenados en el HDFS del clúster.

> **Note**
>
> La consola de Amazon EMR no le permite eliminar un clúster de la vista de lista una vez terminado el clúster. Un clúster terminado desaparece de la consola cuando Amazon EMR borra sus metadatos.

Console

Para terminar el clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Seleccione Clústeres y, a continuación, elija el clúster que desea terminar.
3. En el menú desplegable Acciones, seleccione Terminar clúster.
4. Seleccione Terminar en el cuadro de diálogo. Según la configuración del clúster, la terminación puede tardar de 5 a 10 minutos. Para obtener más información sobre cómo terminar clústeres en Amazon EMR, consulte Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera..

CLI

Para terminar el clúster con el AWS CLI

1. Inicie el proceso de terminación del clúster con el siguiente comando. *<myClusterId>*Sustitúyalo por el ID del clúster de muestra. El comando no devuelve resultados.

```
aws emr terminate-clusters --cluster-ids <myClusterId>
```

2. Para comprobar que el proceso de terminación del clúster esté en curso, compruebe el estado del clúster con el siguiente comando.

```
aws emr describe-cluster --cluster-id <myClusterId>
```

El siguiente ejemplo está en formato JSON. El clúster `Status` debe cambiar de **`TERMINATING`** a **`TERMINATED`**. Según la configuración del clúster, la terminación puede

tardar de 5 a 10 minutos. Para obtener más información sobre la terminación de un clúster de Amazon EMR, consulte Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera..

```
{
    "Cluster": {
        "Id": "j-xxxxxxxxxxxxx",
        "Name": "My Cluster Name",
        "Status": {
            "State": "TERMINATED",
            "StateChangeReason": {
                "Code": "USER_REQUEST",
                "Message": "Terminated by user request"
            }
        }
    }
}
```

Eliminar recursos de S3

Para evitar cargos adicionales, debe eliminar el bucket de Amazon S3. Al eliminar el bucket, se eliminan todos los recursos de Amazon S3 de este tutorial. El bucket debe contener:

- El PySpark guion
- El conjunto de datos de entrada
- La carpeta de resultados de salida
- La carpeta de archivos de registro

Puede que tenga que tomar medidas adicionales para eliminar los archivos almacenados si guardó el PySpark script o el resultado en una ubicación diferente.

> **Note**
>
> Debe terminar el clúster antes de eliminar el bucket. De lo contrario, es posible que no se le permita vaciar el bucket.

Para ello, siga las instrucciones de Eliminar un bucket de S3 en la Guía del usuario de Amazon Simple Storage Service.

Pasos a seguir a continuación

Ya ha lanzado su primer clúster de Amazon EMR de principio a fin. También ha completado tareas esenciales de EMR, como preparar y enviar aplicaciones de macrodatos, ver los resultados y terminar un clúster.

Utilice los siguientes temas para obtener más información sobre cómo personalizar el flujo de trabajo de Amazon EMR.

Explorar las aplicaciones de macrodatos para Amazon EMR

Descubra y compare las aplicaciones de macrodatos que puede instalar en un clúster en la Guía de publicación de Amazon EMR. La guía de versiones detalla cada versión de EMR e incluye consejos para usar marcos como Spark y Hadoop en Amazon EMR.

Planificar el hardware, las redes y la seguridad de los clústeres

En este tutorial, creó un clúster de EMR sencillo sin configurar las opciones avanzadas. Las opciones avanzadas le permiten especificar los tipos de EC2 instancias de Amazon, las redes de clústeres y la seguridad de los clústeres. Para obtener más información sobre la planificación y el lanzamiento de un clúster que cumpla sus requisitos, consulte Planificación, configuración y lanzamiento de clústeres de Amazon EMR y Seguridad en Amazon EMR.

Administrar clústeres

Profundice en el trabajo con clústeres en ejecución en Administración de clústeres de Amazon EMR. Para administrar un clúster, puede conectarse al clúster, depurar los pasos y hacer un seguimiento de las actividades y el estado del clúster. También puede ajustar los recursos del clúster en respuesta a las demandas de carga de trabajo con el escalado administrado por EMR.

Utilizar una interfaz diferente

Además de la consola Amazon EMR, puede gestionar Amazon EMR mediante la API del AWS Command Line Interface servicio web o una de las muchas compatibles. AWS SDKs Para obtener más información, consulte Interfaces de administración.

También puede interactuar con las aplicaciones instaladas en los clústeres de Amazon EMR de muchas maneras. Algunas aplicaciones, como Apache Hadoop, publican interfaces web que puede consultar. Para obtener más información, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

Consultar el blog técnico de EMR

Para ver ejemplos de tutoriales y un análisis técnico detallado sobre las nuevas características de Amazon EMR, consulte el blog de macrodatos de AWS.

Administración de clústeres de Amazon EMR con la consola

La consola ofrece una interfaz actualizada que le proporciona una forma intuitiva de administrar su entorno de Amazon EMR y le brinda un cómodo acceso a la documentación, la información del producto y otros recursos.

Capacidades de la consola

La consola de Amazon EMR está disponible en la siguiente URL:

- URL de la consola: /emr https://console.aws.amazon.com

En la siguiente tabla, se enumera el estado de los principales componentes de la consola de Amazon EMR.

Componente de la consola de Amazon EMR	Consola
EMR Studio	✓
Crear y administrar clústeres	✓
Bloqueo del acceso público	✓
Supervise los CloudWatch eventos de Amazon	✓
Configuraciones de seguridad	✓
Clústeres virtuales (Amazon EMR en EKS)	✓
Ver y gestionar sus subredes de Amazon Virtual Private Cloud[1]	✓
Cuadernos[2]	✓

[1] En la consola, puede ver y administrar las subredes de Amazon VPC en la sección Redes al crear un clúster.

[2] Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Resumen de las diferencias

En esta sección, se describen las capacidades de la experiencia de consola de Amazon EMR. Estas capacidades se dividen en las categorías siguientes:

- Compatibilidad de clústeres en la consola
- Creación de clústeres
- Visualización o edición de los detalles del clúster
- Visualización y búsqueda de clústeres
- Diferencias al trabajar con configuraciones de seguridad

Compatibilidad de clústeres en la consola

En algunos casos, es posible que un clúster que haya creado no sea compatible con la consola. En la siguiente lista se describen los requisitos de compatibilidad de la consola Amazon EMR.

- La consola admite los clústeres creados en Amazon EMR 5.20.1 y versiones posteriores.
- Puede clonar clústeres que utilizan el escalado automático en la consola, pero solo puede crear nuevos clústeres si desea escalarlos manualmente o utilizar el escalado administrado.

Para crear clústeres de la versión 5.20.1 y anteriores y trabajar con ellos, puede utilizar AWS Command Line Interface (AWS CLI) o el AWS SDK.

Creación de clústeres

Funcionalidad	Consola	

Funcionalidad	Consola
Terminología: tipos de nodos de clústeres de Amazon EMR	Principal, secundario, tarea
Versiones compatibles con Amazon EMR[1]	Amazon EMR 5.20.1 y versiones posteriores
Lanzamiento rápido de un clúster	Utilice el botón Crear clúster situado en el panel Resumen. El nombre del clúster no puede contener los caractere s <, >, $, | o ` (acento grave).
Configuración de un tiempo de espera de aprovisionamiento de spot	Defina un periodo de tiempo de espera para aprovisionar instancias para cada flota en el clúster.
Funciones de servicio y función de perfil de EC2 instancia de Amazon	La consola no crea roles predeterminados; debe crear roles con la consola de IAM o seleccionar un rol de IAM ya creado
Visibilidad del clúster	Desde la consola de Amazon EMR, no puede hacer que un clúster sea visible para todos los usuarios; su política de IAM determina el acceso al clúster

Funcionalidad	Consola	
Redes: configurar subredes privadas	Debe configurar los puntos de conexión de Amazon S3 y las puertas de enlace NAT desde sus respectivas consolas Amazon S3 y Amazon VPC	
Vista coherente del sistema de archivos de EMR (EMRFS CV)	Con el lanzamiento de Amazon S3 de gran read-after-write consistencia el 1 de diciembre de 2020, no necesita usar EMRFS CV con sus clústeres de EMR.	
Debugging	Puede depurar los trabajos mediante la interfaz de usuario de la aplicación en la página de detalles del clúster	

[1] No puede crear ni editar clústeres con versiones anteriores a Amazon EMR 5.20.1 en la consola, pero los clústeres existentes creados con versiones anteriores a la 5.20.1 seguirán funcionando. Para crear y editar clústeres con versiones de Amazon EMR anteriores a 5.20.1, utilice la API o la CLI. Puede ver todos los clústeres con la consola, pero es posible que las consolas creadas con versiones anteriores a la versión 5.20.1 no sean compatibles con las funciones más recientes.

Visualización y búsqueda de clústeres

La siguiente tabla muestra cómo puede utilizar la consola Amazon EMR para ver y buscar clústeres.

> **Note**
>
> La aplicación de un filtro de datos a la lista de clústeres consulta toda la base de datos. Sin embargo, al ingresar una cadena de texto en el cuadro de búsqueda, la búsqueda solo se aplica a los resultados que la lista haya cargado en el cliente.

Funcionalidad	Consola	
Ver detalles del clúster	Puede seleccionar el ID del clúster para ver detalles exhaustivos del clúster, como las opciones de configura ción, la aplicación UIs persistente y los registros.	
Búsqueda de clústeres	Utilice un único campo de búsqueda para ingresar consultas de búsqueda de texto y para crear y aplicar filtros de datos como “Estado = Cualquier estado activo”.	
Búsqueda de clústeres con errores	Para buscar clústeres con errores, aplique el filtro Estado = Terminado con errores.	

Visualización o edición de los detalles del clúster

Funcionalidad	Consola	
Ver las instancias de sus grupos de instancias y flotas de instancias, junto con las opciones de escalado, aprovisionamiento, redimensi onamiento y terminación	Consulte las opciones y los detalles de las instancias en la pestaña Instancias. Consulte las opciones de terminación en la pestaña Propiedades.	
	Vea las configuraciones del clúster en la pestaña	

Funcionalidad	Consola	
Visualización de aplicaciones UIs, registros y configuraciones (interfaz de usuario de Apache Spark, servicio Spark History, interfaz de usuario de Apache Tez, servidor de cronología YARN)	Configuraciones. Inicie una interfaz de usuario de aplicación persistente y activa para ver los registros de una aplicación desde la pestaña Aplicaciones.	
Exportación de un clúster a la CLI	La opción está disponible en los menús de acciones de la vista de lista y de detalles del clúster, como “Ver el comando para clonar un clúster”	

Diferencias al trabajar con configuraciones de seguridad

Funcionalidad	Consola	
Clonación de configuraciones de seguridad	✓	
Gobernanza federada mediante Trino y Apache Ranger	✓	
Uso de un rol de tiempo de ejecución para enviar el trabajo a un clúster[1]	✓	

Funcionalidad	Consola	
Autorización del acceso a los datos del sistema de archivos EMR (EMRFS)	Puntos de acceso de Amazon S3	
AWS Lake Formation controles de acceso	Roles de tiempo de ejecución	

[1] Para transferir un rol durante el envío de pasos, el clúster debe usar una configuración de seguridad con una política de permisos de IAM adjunta para que el usuario pueda transferir solo los roles aprobados y sus trabajos puedan acceder a los recursos de Amazon EMR. Para obtener más información, consulte Roles en tiempo de ejecución para los pasos de Amazon EMR.

Amazon EMR Studio

Amazon EMR Studio es un entorno de desarrollo integrado (IDE) basado en web para cuadernos de Jupyter completamente administrados que se ejecutan en clústeres de Amazon EMR. Puede configurar un estudio de EMR para que su equipo desarrolle, visualice y depure aplicaciones escritas en R, Python, Scala y. PySpark EMR Studio está integrado con AWS Identity and Access Management (IAM) e IAM Identity Center para que los usuarios puedan iniciar sesión con sus credenciales corporativas.

Puede crear un EMR Studio sin costo alguno. Cuando utiliza EMR Studio, se aplican cargos por almacenamiento de Amazon S3 y por clústeres de Amazon EMR. Para obtener detalles y aspectos destacados del producto, consulte la página de servicio de Amazon EMR Studio.

Características principales de EMR Studio

Amazon EMR Studio ofrece las siguientes características:

- Autentique a los usuarios con AWS Identity and Access Management (IAM) o AWS IAM Identity Center con o sin una propagación de identidad confiable y su proveedor de identidad empresarial.
- Lance clústeres de Amazon EMR y acceda a ellos bajo demanda para ejecutar trabajos del cuaderno de Jupyter.
- Conéctese a Amazon EMR en los clústeres de EKS para enviar el trabajo como ejecuciones de trabajo.
- Explore y guarde cuadernos de muestra. Para obtener más información sobre cuadernos de ejemplo, consulte el repositorio de ejemplos de GitHub cuadernos de EMR Studio.
- Analice los datos con Python PySpark, Spark Scala, Spark R o SparkSQL e instale kernels y bibliotecas personalizados.
- Colabore en tiempo real con otros usuarios del mismo espacio de trabajo. Para obtener más información, consulte Configuración de la colaboración en el espacio de trabajo en EMR Studio.
- Utilice el Explorador de SQL de EMR Studio para examinar su catálogo de datos, ejecutar consultas SQL y descargar los resultados antes de trabajar con los datos de un cuaderno.
- Ejecute cuadernos parametrizados como parte de los flujos de trabajo programados con una herramienta de orquestación como Apache Airflow o Amazon Managed Workflows para Apache Airflow. Para obtener más información, consulte Organizar trabajos de análisis en Notebooks EMR mediante MWAA en el blog Big Data. AWS

- Enlaza repositorios de código como y. GitHub BitBucket
- Haga un seguimiento y depure las tareas mediante el servidor de historial de Spark, la interfaz de usuario de Tez o el servidor de cronogramas YARN.

EMR Studio cumple con los requisitos de la HIPAA y cuenta con la certificación de HITRUST CSF y de SOC 2. Para obtener más información sobre el cumplimiento de la HIPAA para AWS los servicios, consulte. https://aws.amazon.com/compliance/hipaa-compliance/ Para obtener más información sobre el cumplimiento de los servicios del CSF de HITRUST, consulte AWS . https://aws.amazon.com/compliance/hitrust/

EMR Studio también FedRamp es compatible. Para obtener más información sobre los programas de conformidad que Amazon EMR cumple, consulte Validación de conformidad para Amazon EMR. Para obtener más información sobre los programas de conformidad adicionales para los AWS servicios, consulte los AWS servicios incluidos en el ámbito de aplicación por programa de conformidad.

Entorno de desarrollo integrado Amazon SageMaker Unified Studio

Amazon SageMaker Unified Studio proporciona un entorno de desarrollo integrado (IDE) para sus cuadernos Jupyter que se ejecuta en Amazon EMR en clústeres EC2 o mediante conexiones informáticas EMR Serverless. Al combinar la potencia de Amazon EMR con las capacidades de end-to-end flujo de trabajo de Amazon SageMaker Unified Studio, los equipos pueden optimizar la preparación de datos, el desarrollo de canalizaciones y la experimentación con el aprendizaje automático en un único entorno. Amazon EMR SageMaker revoluciona el procesamiento de macrodatos al admitir marcos de código abierto como Apache Spark, Trino y Apache Flink. Elimine las complejidades de la administración de la infraestructura y, al mismo tiempo, escale las cargas de trabajo de análisis sin esfuerzo. Para obtener más información, consulte Amazon EMR.

Historial de características de Amazon EMR Studio

En esta tabla se enumeran las actualizaciones de la capacidad de Escalado administrado de Amazon EMR.

Fecha de publicación	Funcionalidad
5 de enero de 2024	Se agregó soporte para EMR Studio en AWS GovCloud (EE. UU. Este) y AWS GovCloud (EE. UU. Oeste).

Fecha de publicación	Funcionalidad
26 de noviembre de 2023	Se agregó la compatibilidad con la propagación de identidades de confianza para EMR Studio con la autenticación de IAM Identity Center.
26 de octubre de 2023	Se agregó la posibilidad de crear una aplicación de EMR sin servidor con capacidad interactiva.
28 de febrero de 2023	Se agregó soporte clave AWS KMS administrado por el cliente para el almacenamiento de registros de aplicaciones para aplicaciones EMR Serverless.
23 de febrero de 2023	Se agregó la creación de roles de IAM con un solo clic para el envío de trabajos de EMR sin servidor. Se agregó la búsqueda de ECR para cuando se selecciona una imagen personalizada para las aplicaciones de EMR sin servidor.
27 de enero de 2023	Los cuadernos de ejecución Headless pueden hacer un seguimiento del progreso de la ejecución de cada celda con magia de `%execute_notebook` .
23 de enero de 2023	Las aplicaciones persistentes se han optimizado para acelerar los tiempos de lanzamiento.

Cómo funciona Amazon EMR Studio

Un Amazon EMR Studio es un recurso de Amazon EMR que se crea para un equipo de usuarios. Cada estudio es un entorno de desarrollo integrado basado en web autónomo para cuadernos de Jupyter que se ejecutan en clústeres de Amazon EMR. Los usuarios inician sesión en un estudio con sus credenciales corporativas.

Cada estudio de EMR que cree utiliza los siguientes recursos: AWS

- Una Amazon Virtual Private Cloud (VPC) con subredes: los usuarios ejecutan kernels y aplicaciones de Studio en Amazon EMR y clústeres de Amazon EMR en EKS de la VPC especificada. Un EMR Studio puede conectarse a cualquier clúster de las subredes que especifique al crear el estudio.

- Roles y políticas de permisos de IAM: para administrar los permisos de los usuarios, debe crear políticas de permisos de IAM que se asocian a la identidad de IAM de un usuario o a un rol de usuario. EMR Studio también utiliza un rol de servicio de IAM y grupos de seguridad para interoperar con otros servicios. AWS Para obtener más información, consulte Control de acceso y Definir grupos de seguridad para controlar el tráfico de red de EMR Studio.
- Grupos de seguridad: EMR Studio usa grupos de seguridad para establecer un canal de red seguro entre el estudio y un clúster de EMR.
- Una ubicación de copia de seguridad de Amazon S3: EMR Studio guarda el trabajo del cuaderno en una ubicación de Amazon S3.

Los siguientes pasos describen cómo crear y administrar un EMR Studio:

1. Cree un Studio en su interior Cuenta de AWS con la autenticación de IAM o IAM Identity Center. Para obtener instrucciones, consulte Configurar un EMR Studio.
2. Asigne usuarios o grupos al estudio. Utilice políticas de permisos para establecer permisos específicos para cada usuario. Para obtener más información, consulte el tema Asignar y administrar usuarios de EMR Studio.
3. Comience a monitorear las acciones de EMR Studio con AWS CloudTrail eventos. Para obtener más información, consulte Monitorear las acciones de Amazon EMR Studio.
4. Ofrezca más opciones de clústeres a los usuarios de Studio con plantillas de clúster y puntos de conexión administrados de Amazon EMR en EKS.

Autenticación e inicio de sesión de los usuarios

Amazon EMR Studio admite dos modos de autenticación: el modo de autenticación de IAM y el modo de autenticación de IAM Identity Center. El modo IAM usa AWS Identity and Access Management (IAM), mientras que el modo IAM Identity Center usa. AWS IAM Identity Center Cuando crea un EMR Studio, debe elegir el modo de autenticación para todos los usuarios de ese estudio.

Modo de autenticación de IAM

Con el modo de autenticación de IAM, puede utilizar la autenticación de IAM o la federación de IAM.

La autenticación de IAM le permite administrar las identidades de IAM, como los usuarios, los grupos y los roles de IAM. Concede acceso a los usuarios a un estudio con políticas de permisos de IAM y el control de acceso basado en atributos (ABAC).

La federación de IAM le permite establecer la confianza entre un proveedor de identidades (IdP) externo AWS y, de este modo, puede administrar las identidades de los usuarios a través de su IdP.

Modo de autenticación de IAM Identity Center

El modo de autenticación de IAM Identity Center permite conceder a los usuarios acceso federado a un EMR Studio. Puede usar IAM Identity Center para autenticar usuarios y grupos desde su directorio de IAM Identity Center, su directorio corporativo existente o un IdP externo, como Azure Active Directory (AD). A continuación, debe administrar los usuarios con su proveedor de identidades (IdP).

EMR Studio admite el uso de los siguientes proveedores de identidades para IAM Identity Center:

- AWS Managed Microsoft AD y Active Directory autoadministrado: para obtener más información, consulte Conectarse al directorio de Microsoft AD.
- Proveedores basados en SAML: para ver una lista completa, consulte Proveedores de identidades compatibles.
- El directorio del IAM Identity Center: para obtener más información, consulte Administrar identidades en el IAM Identity Center y Trusted Identity Propagation entre aplicaciones en la Guía del AWS IAM Identity Center usuario.

Cómo afecta la autenticación al inicio de sesión y a la asignación de usuarios

El modo de autenticación que elija para EMR Studio afecta a la forma en que los usuarios inician sesión en un estudio, a la forma en que asigna un usuario a un estudio y a la forma en que autoriza (otorga permisos a) los usuarios para que realicen acciones como la creación de nuevos clústeres de Amazon EMR.

En la siguiente tabla se resumen los métodos de inicio de sesión de EMR Studio según el modo de autenticación.

Opciones de inicio de sesión de EMR Studio por modo de autenticación

Modo de autenticación	Método de inicio de sesión	Descripción
• IAM (autenticación y federación)	URL de EMR Studio	Los usuarios inician sesión en un estudio mediante la URL de acceso al estudio. Por ejemplo, `https://xxxxxxxxxxxxxxxxxx`

Modo de autenticación	Método de inicio de sesión	Descripción
• IAM Identity Center		`xxxx.emrstudio-prod.us-east-1.amazonaws.com` . Los usuarios introducen las credenciales de IAM cuando se utiliza la autenticación de IAM. Cuando utiliza la federación de IAM o IAM Identity Center, EMR Studio redirige a los usuarios a la URL de inicio de sesión del proveedor de identidades para introducir las credenciales. En el contexto de la federación de identidades, esta opción de inicio de sesión se denomina inicio de sesión iniciado por el proveedor de servicios (SP).
• (Federación de) IAM • IAM Identity Center	Portal del proveedor de identidades (IdP)	Los usuarios inician sesión en el portal de su proveedor de identidades, como Azure Portal, e inician la consola de Amazon EMR. Tras lanzar la consola de Amazon EMR, los usuarios seleccionan y abren un estudio de la lista Studios. También puede configurar EMR Studio como una aplicación SAML para que los usuarios puedan iniciar sesión en un estudio específico desde el portal de su proveedor de identidades. Para obtener instrucciones, consulte Para configurar un EMR Studio como una aplicación SAML en el portal de su IdP. En el contexto de la federación de identidades, esta opción de inicio de sesión se denomina inicio de sesión iniciado por el proveedor de identidades (IdP).

Modo de autenticación	Método de inicio de sesión	Descripción
• IAM (autenticación)	AWS Management Console	Los usuarios inician sesión AWS Management Console con las credenciales de IAM y abren un estudio de la lista de estudios de la consola de Amazon EMR.

En la siguiente tabla se describe la asignación y autorización de usuarios en EMR Studio por modo de autenticación.

Asignación y autorización de usuarios de EMR Studio por modo de autenticación

Modo de autenticación	Asignación de usuarios	Autorización de usuarios
IAM (autenticación y federación)	Permita la acción `CreateStudioPresignedUrl` en una política de permisos de IAM asociada a una identidad de IAM (usuario, grupo o rol). En el caso de los usuarios federados, permita la acción `CreateStudioPresignedUrl` en IAM en la política de permisos que configure para el rol de IAM que utilice para la federación. Utilice el control de acceso basado en atributos (ABAC) para especificar el estudio o estudios a los que puede acceder el usuario. Para obtener instrucciones, consulte Asignar un usuario o grupo a un EMR Studio.	Defina políticas de permisos de IAM que permitan determinadas acciones de EMR Studio. En el caso de los usuarios nativos, asocie la política de permisos de IAM a una identidad de IAM (usuario, grupo o rol). En el caso de los usuarios federados, permita las acciones de Studio en la política de permisos que configure para el rol de IAM que utilice para la federación. Para obtener más información, consulte Configurar los permisos de usuario de EMR Studio para Amazon EC2 o Amazon EKS.

Modo de autenticación	Asignación de usuarios	Autorización de usuarios
IAM Identity Center	En el caso de un Studio creado con `IdCUserAssignment` en estado `REQUIRED`, asigne los usuarios al Studio con una política de sesión específica. Para obtener más información, consulte Asignar un usuario o grupo a un EMR Studio. En el caso de un Studio creado con `IdCUserAssignment` en estado `OPTIONAL`, cualquier usuario o grupo de Identity Center puede acceder al Studio.	Opcional: defina políticas de sesión de IAM que permitan determinadas acciones de EMR Studio. Asigne una política de sesión a un usuario al asignar el usuario a un estudio. Para obtener más información, consulte Permisos de usuario para el modo de autenticación de IAM Identity Center.

Control de acceso

En Amazon EMR Studio, debe configurar los permisos de autorización de usuarios con políticas de AWS Identity and Access Management (IAM) basadas en la identidad. Con estas políticas, debe especificar las acciones y los recursos permitidos, así como las condiciones en las que se permiten las acciones.

Permisos de usuario para el modo de autenticación de IAM

Para establecer los permisos de usuario al utilizar la autenticación de IAM en EMR Studio, debe permitir acciones como `elasticmapreduce:RunJobFlow` en una política de permisos de IAM. Puede crear una o más políticas de permisos para utilizarlas. Por ejemplo, puede crear una política básica que no permita a un usuario crear nuevos clústeres de Amazon EMR y otra política que sí permita la creación de clústeres. Para obtener una lista de las acciones de Studio, consulte AWS Identity and Access Management permisos para los usuarios de EMR Studio.

Permisos de usuario para el modo de autenticación de IAM Identity Center

Al utilizar la autenticación de IAM Identity Center, debe crear un único rol de usuario de EMR Studio. El rol de usuario es un rol de IAM dedicado que un estudio asume cuando un usuario inicia sesión.

Asocie las políticas de sesión de IAM al rol de usuario de EMR Studio. Una política de sesión es un tipo especial de política de permisos de IAM que limita lo que un usuario federado puede hacer durante una sesión de inicio de sesión en Studio. Las políticas de sesión le permiten establecer permisos específicos para un usuario o grupo sin crear varios roles de usuario para EMR Studio.

Al asignar usuarios y grupos a un estudio, debe asignar una política de sesión a ese usuario o grupo para aplicar permisos específicos. También puede actualizar la política de sesión de un usuario o grupo en cualquier momento. Amazon EMR almacena cada asignación de políticas de sesión que cree.

Para obtener más información sobre las políticas de sesión, consulte el tema Políticas y permisos en la Guía del usuario de AWS Identity and Access Management .

Workspaces

Los espacios de trabajo son los componentes principales de Amazon EMR Studio. Para organizar los cuadernos, los usuarios crean uno o más espacios de trabajo en un estudio. Para obtener más información, consulte Conozca los espacios de trabajo de EMR Studio.

Al igual que en los espacios de trabajo JupyterLab, un espacio de trabajo conserva el estado del trabajo realizado en un cuaderno. Sin embargo, la interfaz de usuario de Workspace amplía la JupyterLabinterfaz de código abierto con herramientas adicionales que le permiten crear y adjuntar clústeres de EMR, ejecutar trabajos, explorar cuadernos de muestra y vincular repositorios de Git.

En la siguiente lista se incluyen las principales características de los espacios de trabajo de EMR Studio:

- La visibilidad del espacio de trabajo se basa en Studio. Los espacios de trabajo que cree en un estudio no están visibles en otros estudios.
- De forma predeterminada, los espacios de trabajo son compartidos y pueden verlos todos los usuarios de Studio. Sin embargo, solo un usuario puede abrir y trabajar en un espacio de trabajo a la vez. Para trabajar simultáneamente con otros usuarios, puede Configuración de la colaboración en el espacio de trabajo en EMR Studio.
- Puede colaborar simultáneamente con otros usuarios en un espacio de trabajo si activa la colaboración en el espacio de trabajo. Para obtener más información, consulte Configuración de la colaboración en el espacio de trabajo en EMR Studio.
- Los cuadernos de un espacio de trabajo comparten el mismo clúster de EMR para ejecutar comandos. Puede adjuntar un espacio de trabajo a un clúster de Amazon EMR que se ejecute en Amazon EC2 o a un clúster virtual y punto de conexión gestionado de Amazon EMR en EKS.

- Los espacios de trabajo pueden cambiar a otra zona de disponibilidad que asocie a las subredes de un estudio. Puede detener y reiniciar un espacio de trabajo para iniciar el proceso de conmutación por error. Al reiniciar un espacio de trabajo, EMR Studio lanza el espacio de trabajo en una zona de disponibilidad diferente de la VPC del estudio cuando el estudio está configurado con acceso a varias zonas de disponibilidad. Si el estudio solo tiene una zona de disponibilidad, EMR Studio intenta lanzar el espacio de trabajo en una subred diferente. Para obtener más información, consulte Resolver problemas de conectividad con el espacio de trabajo.
- Un espacio de trabajo puede conectarse a clústeres de cualquiera de las subredes asociadas al estudio.

Para obtener más información sobre cómo crear y configurar espacios de trabajo de EMR Studio, consulte Conozca los espacios de trabajo de EMR Studio.

Almacenamiento de cuadernos en Amazon EMR Studio

Cuando utiliza un espacio de trabajo, EMR Studio guarda automáticamente las celdas de los archivos del cuaderno con una cadencia normal en la ubicación de Amazon S3 asociada a su estudio. Este proceso de copia de seguridad conserva el trabajo entre sesiones para que pueda volver a él más adelante sin tener que realizar cambios en un repositorio de Git. Para obtener más información, consulte Guardar contenido del espacio de trabajo en EMR Studio.

Al eliminar un archivo de cuaderno de un espacio de trabajo, EMR Studio elimina automáticamente la versión de copia de seguridad de Amazon S3. Sin embargo, si elimina un espacio de trabajo sin eliminar primero los archivos del cuaderno, los archivos del cuaderno permanecerán en Amazon S3 y seguirán generando gastos de almacenamiento. Para obtener más información, consulte Eliminar un espacio de trabajo y archivos de cuaderno en EMR Studio.

Características, requisitos y límites de EMR Studio

En este tema se incluyen aspectos a tener en cuenta al trabajar con Amazon EMR Studio, como las regiones y las herramientas, los requisitos del clúster y las limitaciones técnicas.

Consideraciones

Tenga en cuenta lo siguiente cuando trabaje con EMR Studio:

- EMR Studio está disponible en las siguientes versiones: Regiones de AWS
  - Este de EE. UU. (Ohio) (us-east-2)

- Este de EE. UU. (Norte de Virginia) (us-east-1)
- EE. UU. Oeste (Norte de California) (us-west-1)
- Oeste de EE. UU. (Oregón) (us-west-2)
- África (Ciudad del Cabo) (af-south-1)
- Asia-Pacífico (Hong Kong) (ap-east-1)
- Asia-Pacífico (Yakarta) (ap-southeast-3)*
- Asia Pacífico (Melbourne) (ap-southeast-4)*
- Asia Pacífico (Bombay) (ap-south-1)
- Asia Pacific (Osaka) (ap-northeast-3)*
- Asia-Pacífico (Seúl) (ap-northeast-2)
- Asia-Pacífico (Singapur) (ap-southeast-1)
- Asia-Pacífico (Sídney) (ap-southeast-2)
- Asia-Pacífico (Tokio) (ap-northeast-1)
- Canadá (centro) (ca-central-1)
- Europa (Fráncfort) (eu-central-1)
- Europa (Irlanda) (eu-west-1)
- Europa (Londres) (eu-west-2)
- UE (Milán) (eu-south-1)
- UE (París) (eu-west-3)
- Europa (España) (eu-south-2)
- Europa (Estocolmo) (eu-north-1)
- Europa (Zúrich) (eu-central-2)*
- Israel (Tel Aviv) (il-central-1)*
- Medio Oriente (EAU) (me-central-1)*
- América del Sur (São Paulo) (sa-east-1)
- AWS GovCloud (US-Este) (gov-us-east-1)
- AWS GovCloud (EEUU-Oeste) (gov-us-west-1)

* La interfaz de usuario de Spark en directo no se admite en estas regiones.

- Para permitir a los usuarios aprovisionar nuevos clústeres de EMR que se ejecuten en Amazon EC2 para un espacio de trabajo, puedes asociar un EMR Studio a un conjunto de plantillas de

clústeres. Los administradores pueden definir plantillas de clúster con Service Catalog y elegir si un usuario o un grupo puede acceder a las plantillas de clúster o a ninguna dentro del estudio.

- Cuando defina los permisos de acceso a los archivos de bloc de notas almacenados en Amazon S3 o de los que lee secretos AWS Secrets Manager, utilice el rol de servicio Amazon EMR. Estos permisos no admiten políticas de sesión.
- Puede crear varios estudios de EMR para controlar el acceso a los clústeres de EMR en diferentes. VPCs
- Úselo AWS CLI para configurar Amazon EMR en los clústeres de EKS. A continuación, puede utilizar la interfaz de Studio para asociar clústeres a los espacios de trabajo con un punto de conexión administrado para ejecutar trabajos de cuaderno.
- Al utilizar la propagación de identidades de confianza con Amazon EMR, hay consideraciones adicionales que también se aplican a EMR Studio. Para obtener más información, consulte Consideraciones y limitaciones de Amazon EMR con la integración de Identity Center.
- EMR Studio no admite los siguientes comandos mágicos de Python:
  - `%alias`
  - `%alias_magic`
  - `%automagic`
  - `%macro`
  - `%%js`
  - `%%javascript`
  - Modificar `proxy_user` mediante `%configure`
  - Modificar `KERNEL_USERNAME` mediante `%env` o `%set_env`
- Los clústeres de Amazon EMR en EKS no admiten SparkMagic comandos para EMR Studio.
- Para escribir instrucciones de Scala de varias líneas en celdas de cuadernos, asegúrese de que todas las líneas, excepto la última, terminen con un punto. En el siguiente ejemplo, se utiliza la sintaxis correcta para las instrucciones de Scala de varias líneas.

```
val df = spark.sql("SELECT * from table_name).
        filter("col1=='value'").
        limit(50)
```

- Para aumentar la seguridad de las aplicaciones fuera de la consola que podría utilizar con Amazon EMR, los dominios de alojamiento de aplicaciones se registran en la lista de sufijos públicos (PSL). Algunos ejemplos de estos dominios de alojamiento son los siguientes: `emrstudio-prod.us-`

`east-1.amazonaws.com`, `emrnotebooks-prod.us-east-1.amazonaws.com`, `emrappui-prod.us-east-1.amazonaws.com`. Para mayor seguridad, si necesita configurar cookies confidenciales en el nombre de dominio predeterminado, le recomendamos que utilice cookies con el prefijo `__Host-`. Esta práctica lo ayuda a proteger su dominio de los intentos de falsificación de solicitudes entre sitios (CSRF). Para obtener más información, consulte la página Set-Cookie en la Red de desarrolladores de Mozilla.

- Los espacios de trabajo y los puntos de enlace de la interfaz de usuario persistente de Amazon EMR Studio utilizan módulos criptográficos validados por FIPS 140 encryption-in-transit, lo que facilita la adopción del servicio para cargas de trabajo reguladas. Para obtener más información sobre los puntos de enlace de la interfaz de usuario persistente, consulte Ver las interfaces de usuario de aplicaciones persistentes en Amazon EMR. Para obtener más información sobre las libretas, consulte Información general sobre Amazon EMR Notebooks.

Problemas conocidos

- Un EMR Studio que utiliza IAM Identity Center con la propagación de identidades de confianza habilitada solo puede asociarse con clústeres de EMR que también utilizan la propagación de identidades de confianza.
- Asegúrese de desactivar las herramientas de administración de proxy como FoxyProxy o SwitchyOmega en el navegador antes de crear un Studio. Los proxies activos pueden provocar errores al seleccionar Crear estudio y pueden generar un mensaje de error de red.
- Los kernels que se ejecutan en clústeres de Amazon EMR en EKS pueden no iniciarse debido a problemas de tiempo de espera. Si se produce un error o un problema al iniciar el kernel, cierre el archivo del cuaderno, apague el kernel y vuelva a abrir el archivo del cuaderno.
- La operación Reiniciar el kernel no funciona según lo esperado cuando se utiliza un clúster de Amazon EMR en EKS. Tras seleccionar Reiniciar el kernel, actualice el espacio de trabajo para que el reinicio surta efecto.
- Si un espacio de trabajo no está conectado a un clúster, aparece un mensaje de error cuando un usuario de Studio abre un archivo de cuaderno e intenta seleccionar un kernel. Para ignorar este mensaje de error, pulse Aceptar, pero debe asociar el espacio de trabajo a un clúster y seleccionar un kernel para poder ejecutar el código del cuaderno.
- Cuando utiliza Amazon EMR 6.2.0 con una configuración de seguridad para configurar la seguridad del clúster, la interfaz del espacio de trabajo aparece en blanco y no funciona como se esperaba. Le recomendamos que utilice una versión compatible diferente de Amazon EMR si desea configurar el cifrado de datos o la autorización de Amazon S3 para EMRFS en un clúster.

EMR Studio funciona con las versiones 5.32.0 (serie 5.x de Amazon EMR) o 6.2.0 (serie 6.x de Amazon EMR) y posteriores de Amazon EMR.

- Cuando Depure Amazon EMR que se ejecuta en Amazon Jobs EC2 , es posible que los enlaces a la interfaz de usuario de Spark en el clúster no funcionen o no aparezcan. Para regenerar los enlaces, cree una nueva celda del cuaderno y ejecute el comando `%%info`.
- Jupyter Enterprise Gateway no limpia los kernels inactivos del nodo principal de un clúster en las siguientes versiones de Amazon EMR: 5.32.0, 5.33.0, 6.2.0 y 6.3.0. Los kernels inactivos consumen recursos de computación y pueden provocar errores en los clústeres que se ejecutan durante mucho tiempo. Puede configurar la limpieza de kernels inactivos de Jupyter Enterprise Gateway mediante el siguiente script de ejemplo. Puede Conexión al nodo principal del clúster de Amazon EMR mediante SSH o enviar el script como un paso. Para obtener más información, consulte Ejecutar comandos y scripts en un clúster de Amazon EMR.

```
#!/bin/bash
sudo tee -a /emr/notebook-env/conf/jupyter_enterprise_gateway_config.py << EOF
c.MappingKernelManager.cull_connected = True
c.MappingKernelManager.cull_idle_timeout = 10800
c.MappingKernelManager.cull_interval = 300
EOF
sudo systemctl daemon-reload
sudo systemctl restart jupyter_enterprise_gateway
```

- Cuando utiliza una política de terminación automática con las versiones 5.32.0, 5.33.0, 6.2.0 o 6.3.0 de Amazon EMR, Amazon EMR marca un clúster como inactivo y puede terminarlo automáticamente incluso si tiene un kernel de Python3 activo. Esto se debe a que al ejecutar un kernel de Python3 no se envía ningún trabajo de Spark al clúster. Para utilizar la terminación automática con un kernel de Python3, le recomendamos que utilice la versión 6.4.0 o posterior de Amazon EMR. Para obtener más información sobre la terminación automática, consulte Uso de una política de finalización automática para la limpieza de clústeres de Amazon EMR..
- Al mostrar un Spark DataFrame en una tabla, las tablas muy anchas pueden quedar truncadas. `%%display` Puede hacer clic con el botón derecho en la salida y seleccionar Crear nueva vista para la salida para obtener una vista desplazable de la salida.
- Al iniciar un núcleo basado en Spark, como PySpark Spark o SparkR, se inicia una sesión de Spark y, al ejecutar una celda en un cuaderno, se ponen en cola los trabajos de Spark de esa sesión. Cuando interrumpes una celda en ejecución, el trabajo de Spark continúa ejecutándose. Para detener el trabajo de Spark, debe usar la interfaz de usuario de Spark en el clúster. Para

obtener instrucciones sobre cómo conectarse a la interfaz de usuario de Spark, consulte Depurar aplicaciones y trabajos con EMR Studio.

- El uso de Amazon EMR Studio Workspaces como usuario root en un Cuenta de AWS provoca un error. `403: Forbidden` Esto se debe a que la configuración de Jupyter Enterprise Gateway en Amazon EMR no permite el acceso al usuario raíz. Recomendamos que no utilice el usuario raíz para sus tareas diarias. Para ver otras opciones de autenticación, consulte AWS Identity and Access Management para Amazon EMR.

Limitaciones de características

Amazon EMR Studio no admite las siguientes características de Amazon EMR:

- Asociar y ejecutar trabajos en clústeres de EMR con una configuración de seguridad que especifique la autenticación de Kerberos
- Clústeres con varios nodos principales
- Clústeres que utilizan EC2 instancias de Amazon basadas en AWS Graviton2 para las versiones 6.x de Amazon EMR anteriores a la 6.9.0 y 5.x inferiores a la 5.36.1

Un Studio que utiliza la propagación de identidades de confianza no admite las siguientes características:

- Creación de clústeres de EMR sin plantilla.
- Utilización de aplicaciones de EMR sin servidor.
- Lanzamiento de clústeres de Amazon EMR en EKS.
- Utilización de un rol de tiempo de ejecución.
- Habilitación de la colaboración en SQL Explorer o Workspace.

Límites de servicio para EMR Studio

En la siguiente tabla se muestran los límites de servicio de EMR Studio.

Elemento	Límite
Estudios de EMR Studio	AWS Máximo de 100 por cuenta

Elemento	Límite
Subredes	Máximo de 5 asociados a cada EMR Studio
Grupos de IAM Identity Center	Máximo de 5 asignados a cada EMR Studio
Usuarios de IAM Identity Center	Máximo de 100 asignados a cada EMR Studio

Prácticas recomendadas sobre VPC y subredes para EMR Studio

Utilice las siguientes prácticas recomendadas para configurar una Amazon Virtual Private Cloud (Amazon VPC) con subredes para EMR Studio:

- Puede especificar un máximo de cinco subredes en la VPC para asociarlas al estudio. Recomendamos que proporcione varias subredes en diferentes zonas de disponibilidad para respaldar la disponibilidad del espacio de trabajo y permitir que los usuarios de Studio accedan a los clústeres de las distintas zonas de disponibilidad. Para obtener más información sobre cómo trabajar con VPCs subredes y zonas de disponibilidad, consulte las subredes en la VPCs Guía del usuario.Amazon Virtual Private Cloud
- Las subredes que especifique deben poder comunicarse entre sí.
- Para permitir que los usuarios vinculen un espacio de trabajo a repositorios de Git alojados públicamente, debe especificar solo las subredes privadas que tengan acceso a internet a través de la traducción de direcciones de red (NAT). Para obtener más información sobre la configuración de una subred privada para Amazon EMR, consulte Subredes privadas.
- Cuando utiliza Amazon EMR en EKS con EMR Studio, debe haber al menos una subred en común entre su estudio y el clúster de Amazon EKS que utiliza para registrar un clúster virtual. De lo contrario, el punto de conexión administrado no aparecerá como opción en los espacios de trabajo de Studio. Puede crear un clúster de Amazon EKS y asociarlo a una subred que pertenezca al estudio, o bien crear un estudio y especificar las subredes del clúster de EKS.
- Si planea usar Amazon EMR en EKS con EMR Studio, seleccione la misma VPC como los nodos de trabajo del clúster de Amazon EKS.

Requisitos del clúster de Amazon EMR

Clústeres de Amazon EMR que se ejecutan en Amazon EC2

Todos los clústeres de Amazon EMR que se ejecuten en Amazon y EC2 que cree para un espacio de trabajo de EMR Studio deben cumplir los siguientes requisitos. Los clústeres que cree mediante la interfaz de EMR Studio cumplen automáticamente estos requisitos.

- El clúster debe usar las versiones 5.32.0 (serie Amazon EMR 5.x) o 6.2.0 (serie Amazon EMR 6.x) o posteriores de Amazon EMR. Puede crear un clúster mediante la consola o el SDK de Amazon EMR y AWS Command Line Interface, a continuación, adjuntarlo a un espacio de trabajo de EMR Studio. Los usuarios del estudio también pueden aprovisionar y asociar clústeres al crear o trabajar en un espacio de trabajo de Amazon EMR. Para obtener más información, consulte Asociar computación a un espacio de trabajo de EMR Studio.
- El clúster debe estar dentro de una Amazon Virtual Private Cloud. La plataforma EC2 -Classic no es compatible.
- El clúster debe tener Spark, Livy y Jupyter Enterprise Gateway instalados. Si planea usar el clúster para el Explorador de SQL, debe instalar Presto y Spark.
- Para usar el Explorador de SQL, el clúster debe usar la versión 5.34.0 o posterior o la versión 6.4.0 o posterior de Amazon EMR y tener instalado Presto. Si quiere especificar el catálogo de datos de AWS Glue como el metaalmacén de Hive para Presto, debe configurarlo en el clúster. Para obtener más información, consulte Uso de Presto con el Catálogo de datos de AWS Glue.
- El clúster debe estar en una subred privada con traducción de direcciones de red (NAT) para usar repositorios de Git alojados públicamente con EMR Studio.

Recomendamos las siguientes configuraciones de clúster cuando trabaje con EMR Studio.

- Configure el modo de despliegue de las sesiones de Spark en el modo de clúster. El modo de clúster coloca los procesos maestros de la aplicación en los nodos básicos y no en el nodo principal de un clúster. De este modo, se alivian las posibles presiones de memoria del nodo principal. Para obtener más información, consulte Cluster Mode Overview en la documentación de Apache Spark.
- Cambie el tiempo de espera de Livy del valor predeterminado de una hora a seis horas, como en el siguiente ejemplo de configuración.

```
{
    "classification":"livy-conf",
        "Properties":{
            "livy.server.session.timeout":"6h",
            "livy.spark.deploy-mode":"cluster"
        }
```

```
}
```

- Cree diversas flotas de instancias con hasta 30 instancias y seleccione varios tipos de instancias en su flota de instancias de spot. Por ejemplo, puede especificar los siguientes tipos de instancias optimizadas para memoria para las cargas de trabajo de Spark: r5.2x, r5.4x, r5.8x, r5.12x, r5.16x, r4.2x, r4.4x, r4.8x, r4.12, etc. Para obtener más información, consulte Planificación y configuración de flotas de instancias para su clúster de Amazon EMR.
- Utilice la estrategia de asignación de capacidad optimizada para las instancias puntuales para ayudar a Amazon EMR a seleccionar instancias de forma eficaz en función de los conocimientos de capacidad en tiempo real de Amazon. EC2 Para obtener más información, consulte Estrategia de asignación para flotas de instancias.
- Habilite el escalamiento administrado en su clúster. Establezca el parámetro de número máximo de nodos principales en la capacidad persistente mínima que planea utilizar y configure el escalamiento en función de una flota de tareas bien diversificada que se ejecute en instancias de spot para ahorrar costos. Para obtener más información, consulte Uso del escalado administrado en Amazon EMR.

También le instamos a que mantenga activado el Bloqueo de acceso público de Amazon EMR y que restrinja el tráfico SSH entrante a orígenes de confianza. El acceso entrante a un clúster permite a los usuarios ejecutar cuadernos en el clúster. Para obtener más información, consulte Uso de Bloquear el acceso público de Amazon EMR y Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

Clústeres de Amazon EMR en EKS

Además de los clústeres de EMR que se ejecutan en Amazon EC2, puede configurar y gestionar Amazon EMR en clústeres de EKS para EMR Studio mediante. AWS CLI Configure los clústeres de Amazon EMR en EKS siguiendo las pautas que se indican a continuación:

- Cree un punto de conexión HTTPS administrado para el clúster de Amazon EMR en EKS. Los usuarios asocian un espacio de trabajo a un punto de conexión administrado. El clúster de Amazon Elastic Kubernetes Service (EKS) que utilice para registrar un clúster virtual debe tener una subred privada para admitir los puntos de conexión administrados.
- Utilice un clúster de Amazon EKS con al menos una subred privada y una traducción de direcciones de red (NAT) cuando desee utilizar repositorios de Git alojados públicamente.
- Evite utilizar Amazon Linux Arm optimizado para Amazon EKS AMIs, que no son compatibles con Amazon EMR en los puntos de conexión gestionados por EKS.

- Evite utilizar AWS Fargate clústeres exclusivos de Amazon EKS, que no son compatibles.

Configuración de Amazon EMR Studio

Esta sección es para los administradores de EMR Studio. En ella se explica cómo configurar un EMR Studio para su equipo y se proporcionan instrucciones para tareas como la asignación de usuarios y grupos, la configuración de plantillas de clústeres y la optimización de Apache Spark para EMR Studio.

Temas

- Permisos de administrador para crear y administrar un EMR Studio
- Configurar un EMR Studio
- Supervise, actualice y elimine los recursos de Amazon EMR Studio
- Cifrado de cuadernos y archivos del espacio de trabajo de EMR Studio
- Definir grupos de seguridad para controlar el tráfico de red de EMR Studio
- Cree AWS CloudFormation plantillas para Amazon EMR Studio
- Establecer el acceso y los permisos para los repositorios basados en Git
- Optimizar los trabajos de Spark en EMR Studio

Permisos de administrador para crear y administrar un EMR Studio

Los permisos de IAM descritos en esta página le permiten crear y administrar un EMR Studio. Para obtener información detallada sobre cada permiso necesario, consulte la Permisos necesarios para administrar un EMR Studio.

Permisos necesarios para administrar un EMR Studio

En la siguiente tabla se enumeran las operaciones relacionadas con la creación y administración de un EMR Studio. En la tabla también se muestran los permisos necesarios para cada operación.

> **Note**
>
> Solo necesita las acciones `SessionMapping` de IAM Identity Center y Studio cuando utiliza el modo de autenticación del IAM Identity Center.

Permisos para crear y administrar un EMR Studio

Operación	Permisos
Crear un estudio	`"elasticmapreduce:CreateStudio", "sso:CreateApplication", "sso:PutApplicationAuthenticationMethod", "sso:PutApplicationGrant", "sso:PutApplicationAccessScope", "sso:PutApplicationAssignmentConfiguration", "iam:PassRole"`
Describir un estudio	`"elasticmapreduce:DescribeStudio", "sso:GetManagedApplicationInstance"`
Enumerar los estudios	`"elasticmapreduce:ListStudios"`
Eliminar un estudio	`"elasticmapreduce:DeleteStudio", "sso:DeleteApplication", "sso:DeleteApplicationAuthenticationMethod", "sso:DeleteApplicationAccessScope", "sso:DeleteApplicationGrant"`
Additional permissions required when you use IAM Identity Center mode	
Asignar usuarios o grupos a un estudio	`"elasticmapreduce:CreateStudioSessionMapping", "sso:GetProfile", "sso:ListDirectoryAssociations", "sso:ListProfiles", "sso:AssociateProfile", "sso-directory:SearchUsers", "sso-directory:SearchGroups", "sso-directory:DescribeUser", "sso-directory:DescribeGroup", "sso:ListInstances", "sso:CreateApplicationAssignment",`

Operación	Permisos
	`"sso:DescribeInstance", "organizations:DescribeOrganization", "organizations:ListDelegatedAdministrators", "sso:CreateInstance", "sso:DescribeRegisteredRegions", "sso:GetSharedSsoConfiguration", "iam:ListPolicies"`
Recuperar los detalles de las tareas de Studio para un usuario o grupo específico	`"sso-directory:SearchUsers", "sso-directory:SearchGroups", "sso-directory:DescribeUser", "sso-directory:DescribeGroup", "sso:DescribeApplication", "elasticmapreduce:GetStudioSessionMapping"`
Enumerar todos los usuarios y grupos asignados a un estudio	`"elasticmapreduce:ListStudioSessionMappings"`
Actualizar la política de sesión asociada a un usuario o grupo asignado a un estudio	`"sso-directory:SearchUsers", "sso-directory:SearchGroups", "sso-directory:DescribeUser", "sso-directory:DescribeGroup", "sso:DescribeApplication", "sso:DescribeInstance", "elasticmapreduce:UpdateStudioSessionMapping"`

Operación	Permisos
Eliminar un usuario o grupo de un estudio	`"elasticmapreduce:DeleteStudioSessionMapping", "sso-directory:SearchUsers", "sso-directory:SearchGroups", "sso-directory:DescribeUser", "sso-directory:DescribeGroup", "sso:ListDirectoryAssociations", "sso:GetProfile", "sso:DescribeApplication", "sso:DescribeInstance", "sso:ListProfiles", "sso:DisassociateProfile", "sso:DeleteApplicationAssignment", "sso:ListApplicationAssignments"`

Para crear una política con permisos de administrador para EMR Studio

1. Siga las instrucciones de Creación de políticas de IAM para crear una política mediante uno de los siguientes ejemplos. Los permisos que necesita dependen del modo de autenticación de EMR Studio.

   Introduzca sus propios valores para los siguientes elementos:

   - Reemplace *<your-resource-ARN>* para especificar el nombre de recurso de Amazon (ARN) del objeto o los objetos que cubre la declaración para sus casos de uso.
   - *<region>*Sustitúyalo por el código del Región de AWS lugar donde planea crear el estudio.
   - *<aws-account_id>*Sustitúyalo por el ID de la AWS cuenta del estudio.
   - Sustituya *<EMRStudio-Service-Role>* y por *<EMRStudio-User-Role>* los nombres de su función de servicio de EMR Studio y de su función de usuario de EMR Studio.

Example Política de ejemplo: permisos de administrador cuando utiliza el modo de autenticación de IAM

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:studio/*"
      ],
      "Action": [
        "elasticmapreduce:CreateStudio",
        "elasticmapreduce:DescribeStudio",
        "elasticmapreduce:DeleteStudio"
      ],
      "Sid": "AllowELASTICMAPREDUCECreatestudio"
    },
    {
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Action": [
        "elasticmapreduce:ListStudios"
      ],
      "Sid": "AllowELASTICMAPREDUCEListstudios"
    },
    {
      "Effect": "Allow",
      "Resource": [
        "arn:aws:iam::123456789012:role/EMRStudioServiceRole"
      ],
      "Action": [
        "iam:PassRole"
      ],
      "Sid": "AllowIAMPassrole"
    }
  ]
```

```
}
```

Example Política de ejemplo: permisos de administrador cuando utiliza el modo de autenticación de IAM Identity Center

> **Note**
>
> Identity Center y el directorio de Identity Center APIs no admiten la especificación de un ARN en el elemento de recurso de una declaración de política de IAM. Para permitir el acceso a IAM Identity Center e IAM Identity Center Directory, los siguientes permisos especifican todos los recursos, “Resource”:“*”, para las acciones de IAM Identity Center. Para obtener información, consulte Acciones, recursos y claves de condición de IAM Identity Center Directory.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:studio/*"
      ],
      "Action": [
        "elasticmapreduce:CreateStudio",
        "elasticmapreduce:DescribeStudio",
        "elasticmapreduce:DeleteStudio",
        "elasticmapreduce:CreateStudioSessionMapping",
        "elasticmapreduce:GetStudioSessionMapping",
        "elasticmapreduce:UpdateStudioSessionMapping",
        "elasticmapreduce:DeleteStudioSessionMapping"
      ],
      "Sid": "AllowELASTICMAPREDUCECreatestudio"
    },
    {
      "Effect": "Allow",
      "Resource": [
```

```
        "*"
      ],
      "Action": [
        "elasticmapreduce:ListStudios",
        "elasticmapreduce:ListStudioSessionMappings"
      ],
      "Sid": "AllowELASTICMAPREDUCEListstudios"
    },
    {
      "Effect": "Allow",
      "Resource": [
        "arn:aws:iam::123456789012:role/EMRStudio-SvcRole",
        "arn:aws:iam::123456789012:role/EMRStudio-User-Role"
      ],
      "Action": [
        "iam:PassRole"
      ],
      "Sid": "AllowIAMPassrole"
    },
    {
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Action": [
        "sso:CreateApplication",
        "sso:PutApplicationAuthenticationMethod",
        "sso:PutApplicationGrant",
        "sso:PutApplicationAccessScope",
        "sso:PutApplicationAssignmentConfiguration",
        "sso:DescribeApplication",
        "sso:DeleteApplication",
        "sso:DeleteApplicationAuthenticationMethod",
        "sso:DeleteApplicationAccessScope",
        "sso:DeleteApplicationGrant",
        "sso:ListInstances",
        "sso:CreateApplicationAssignment",
        "sso:DeleteApplicationAssignment",
        "sso:ListApplicationAssignments",
        "sso:DescribeInstance",
        "sso:AssociateProfile",
        "sso:DisassociateProfile",
        "sso:GetProfile",
        "sso:ListDirectoryAssociations",
```

```
        "sso:ListProfiles",
        "sso-directory:SearchUsers",
        "sso-directory:SearchGroups",
        "sso-directory:DescribeUser",
        "sso-directory:DescribeGroup",
        "organizations:DescribeOrganization",
        "organizations:ListDelegatedAdministrators",
        "sso:CreateInstance",
        "sso:DescribeRegisteredRegions",
        "sso:GetSharedSsoConfiguration",
        "iam:ListPolicies"
      ],
      "Sid": "AllowSSOCreateapplication"
    }
  ]
}
```

2. Asocie la política a su identidad de IAM (usuario, rol o grupo). Para ver cómo hacerlo, consulte Agregar y eliminar permisos de identidad de IAM.

Configurar un EMR Studio

Complete los siguientes pasos para configurar un EMR Studio.

Antes de comenzar

> **Note**
>
> Si planea usar EMR Studio con Amazon EMR en EKS, le recomendamos que primero configure Amazon EMR en EKS para EMR Studio antes de configurar un estudio.

Antes de configurar un EMR Studio, asegúrese de tener los siguientes elementos:

- Un. Cuenta de AWS Para obtener instrucciones, consulte Antes de configurar Amazon EMR.
- Permisos para crear y administrar un EMR Studio. Para obtener más información, consulte the section called “Permisos de administrador para crear un EMR Studio”.
- Un bucket de Amazon S3 en el que EMR Studio puede hacer copias de seguridad de los espacios de trabajo y los archivos de cuadernos de su estudio. Para obtener instrucciones, consulte Crear un bucket en la Guía del usuario de Amazon Simple Storage Service (S3).

- Si desea conectarse a un clúster de Amazon EMR o EC2 Amazon EMR de EKS, o utilizar repositorios de Git, necesitará una Amazon Virtual Private Cloud (VPC) para el estudio y un máximo de cinco subredes. No necesita una VPC para usar EMR Studio con EMR sin servidor. Para obtener consejos sobre cómo configurar las redes, consulte Prácticas recomendadas sobre VPC y subredes para EMR Studio.

Para configurar un EMR Studio

1. Elija un modo de autenticación para Amazon EMR Studio
2. Cree los siguientes recursos de Studio.
   - Crear un rol de servicio de EMR Studio
   - Configurar los permisos de usuario de EMR Studio para Amazon o EC2 Amazon EKS
   - (Opcional) Definir grupos de seguridad para controlar el tráfico de red de EMR Studio.
3. Crear un EMR Studio
4. Asignar un usuario o grupo a un EMR Studio

Una vez que haya completado los pasos de configuración, podrá Uso de un Amazon EMR Studio.

Elija un modo de autenticación para Amazon EMR Studio

EMR Studio admite dos modos de autenticación: el modo de autenticación de IAM y el modo de autenticación de IAM Identity Center. El modo IAM usa AWS Identity and Access Management (IAM), mientras que el modo IAM Identity Center usa. AWS IAM Identity Center Cuando crea un EMR Studio, debe elegir el modo de autenticación para todos los usuarios de ese estudio. Para obtener más información sobre los distintos modos de autenticación, consulte Autenticación e inicio de sesión de los usuarios.

Utilice la siguiente tabla para elegir un modo de autenticación para EMR Studio.

En caso de que...	Recomendamos...
Ya esté familiarizado con la autenticación o la federación de IAM o la haya configurado anteriormente	Modo de autenticación de IAM, que ofrece los siguientes beneficios:

En caso de que...	Recomendamos...
	• Proporciona una configuración rápida para EMR Studio si ya administra identidades como usuarios y grupos en IAM. • Funciona con proveedores de identidad es que son compatibles con OpenID Connect (OIDC) o Security Assertion Markup Language 2.0 (SAML 2.0). • Admite el uso de varios proveedores de identidad con la misma Cuenta de AWS. • Disponible en un amplio número de. Regiones de AWS • Cumple con SOC 2.
AWS ¿Es nuevo en Amazon EMR?	Modo de autenticación de IAM Identity Center, que ofrece las siguientes características: • Permite una fácil asignación de AWS recursos por parte de usuarios y grupos. • Funciona con los proveedores de identidades Microsoft Active Directory y SAML 2.0. • Facilita la configuración de la federació n de varias cuentas para que no tenga que configurar la federación por separado para cada miembro Cuenta de AWS de la organización.

Configuración del modo de autenticación de IAM para Amazon EMR Studio

Con el modo de autenticación de IAM, puede utilizar la autenticación de IAM o la federación de IAM. La autenticación de IAM le permite administrar las identidades de IAM, como los usuarios, los grupos y los roles de IAM. Concede acceso a los usuarios a un estudio con políticas de permisos de IAM y el control de acceso basado en atributos (ABAC). La federación de IAM le permite establecer la confianza entre un proveedor de identidades (IdP) externo AWS y, de este modo, puede administrar las identidades de los usuarios a través de su IdP.

> **Note**
>
> Si ya utiliza IAM para controlar el acceso a AWS los recursos, o si ya ha configurado su proveedor de identidad (IdP) para IAM, Permisos de usuario para el modo de autenticación de IAM consulte para establecer los permisos de usuario cuando utilice el modo de autenticación de IAM para EMR Studio.

Uso de la federación de IAM para Amazon EMR Studio

Para utilizar la federación de IAM para EMR Studio, debe crear una relación de confianza entre usted y Cuenta de AWS su proveedor de identidad (IdP) y permitir que los usuarios federados accedan a. AWS Management Console Los pasos que debe seguir para crear esta relación de confianza varían según el estándar de federación de su IdP.

En general, debe completar las siguientes tareas para configurar la federación con un IdP externo. Para obtener instrucciones completas, consulte Habilitar usuarios federados de SAML 2.0 para acceder a la AWS Management Console y Habilitar el acceso de un agente de identidades personalizado a la AWS Management Console en la Guía del usuario de AWS Identity and Access Management .

1. Recopile información de su IdP. Por lo general, esto implica generar un documento de metadatos para validar las solicitudes de autenticación SAML de su IdP.
2. Cree una entidad de IAM del proveedor de identidades para almacenar información sobre su IdP. Para obtener instrucciones, consulte Creación de proveedores de identidades de IAM.
3. Cree uno o varios roles de IAM para su IdP. EMR Studio asigna un rol a un usuario federado al iniciar sesión. El rol permite al proveedor de identidades solicitar credenciales de seguridad temporales para obtener acceso a AWS. Para obtener instrucciones, consulte Creación de un rol para un proveedor de identidades de terceros (federación). Las políticas de permisos que se asignan al rol determinan lo que los usuarios federados pueden hacer en un estudio de EMR AWS y dentro de él. Para obtener más información, consulte Permisos de usuario para el modo de autenticación de IAM.
4. (Para los proveedores de SAML) Complete la confianza de SAML configurando su IdP con la información AWS y las funciones que desea que asuman los usuarios federados. Este proceso de configuración crea confianza entre la parte de confianza entre su IdP y. AWS Para obtener más información, consulte Configuración de una relación de confianza para usuario autenticado y agregación de notificaciones en el proveedor de identidades SAML 2.0.

Para configurar un EMR Studio como una aplicación SAML en el portal de su IdP

Puede configurar un EMR Studio concreto como aplicación SAML mediante un enlace directo al estudio. Esto permite que los usuarios inicien sesión en su portal de IdP y lancen un estudio específico en lugar de tener que navegar por la consola de Amazon EMR.

- Utilice el siguiente formato para configurar un enlace directo a su EMR Studio como URL de destino tras la verificación de la aserción de SAML.

```
https://console.aws.amazon.com/emr/home?region=<aws-region>#studio/<your-studio-id>/start
```

Configuración del modo de autenticación de IAM Identity Center para Amazon EMR Studio

Para prepararse AWS IAM Identity Center para EMR Studio, debe configurar su fuente de identidad y aprovisionar usuarios y grupos. El aprovisionamiento es el proceso de poner la información de usuarios y grupos a disposición de IAM Identity Center y de las aplicaciones que utilizan IAM Identity Center. Para obtener más información, consulte Aprovisionamiento de usuarios y grupos.

EMR Studio admite el uso de los siguientes proveedores de identidades para IAM Identity Center:

- AWS Managed Microsoft AD y Active Directory autoadministrado: para obtener más información, consulte Conectarse al directorio de Microsoft AD.
- Proveedores basados en SAML: para ver una lista completa, consulte Proveedores de identidades compatibles.
- El directorio de IAM Identity Center: para obtener más información, consulte Administrar identidades en IAM Identity Center.

Para configurar IAM Identity Center para EMR Studio

1. Para configurar IAM Identity Center para EMR Studio, necesita lo siguiente:
   - Una cuenta de administración en su AWS organización si usa varias cuentas en su organización.

> **Note**
>
> Solo debe usar su cuenta de administración para habilitar IAM Identity Center y aprovisionar usuarios y grupos. Tras configurar IAM Identity Center, utilice una cuenta de miembro para crear un EMR Studio y asignar usuarios y grupos. Para obtener más información sobre AWS la terminología, consulte AWS Organizations Terminología y conceptos.

- Si activó el Centro de identidad de IAM antes del 25 de noviembre de 2019, es posible que deba habilitar las aplicaciones que utilizan el Centro de identidad de IAM para las cuentas de su AWS organización. Para obtener más información, consulte Habilitar las aplicaciones integradas en el IAM Identity Center en las cuentas. AWS
- Asegúrese de que los requisitos previos figuren en la página Requisitos previos de IAM Identity Center.

2. Siga las instrucciones de Activar el centro de identidad de IAM para activar el centro de identidad de IAM en el Región de AWS lugar donde desee crear el EMR Studio.
3. Conecte IAM Identity Center con su proveedor de identidades y aprovisione los usuarios y grupos que desee asignar al estudio.

Si usa…	Haga lo siguiente...
Un directorio de Microsoft AD	1. Siga las instrucciones de Conectarse al directorio de Microsoft AD para conectar su Active Directory o AWS Managed Microsoft AD directorio autogestionado mediante AWS Directory Service. 2. Para aprovisionar usuarios y grupos para IAM Identity Center, puede sincronizar los datos de identidades de su AD de origen con IAM Identity Center. Puede sincroniz ar las identidades de su AD de origen de muchas maneras. Una forma es asignar usuarios o grupos de AD a una cuenta de AWS de su organización. Para obtener

Si usa…	Haga lo siguiente...
	instrucciones, consulte Inicio de sesión único. La sincronización puede tardar hasta dos horas. Una vez que complete este paso, los usuarios y grupos sincronizados aparecerán en su almacén de identidades. **Note** Los usuarios y los grupos no aparecen en tu almacén de identidades hasta que no sincronic es la información de usuarios y grupos o utilices el aprovisio namiento de usuarios just-in-time (JIT). Para obtener más informaci ón, consulte Aprovisionamiento cuando los usuarios provienen de Active Directory. 3. (Opcional) Tras sincronizar los usuarios y grupos de AD, puedes eliminar su acceso a la AWS cuenta que configuraste en el paso anterior. Para obtener instrucci ones, consulte Eliminar el acceso de los usuarios.
Un proveedor de identidades externo	Siga las instrucciones de Conectarse a su proveedor de identidades externo.
El directorio de IAM Identity Center	Al crear usuarios y grupos en IAM Identity Center, el aprovisionamiento es automátic o. Para obtener más información, consulte Administrar identidades en el IAM Identity Center.

Ahora puede asignar usuarios y grupos de su almacén de identidades a un EMR Studio. Para obtener instrucciones, consulte Asignar un usuario o grupo a un EMR Studio.

Crear un rol de servicio de EMR Studio

Acerca del rol de servicio de EMR Studio

Cada EMR Studio utiliza un rol de IAM con permisos que permiten al estudio interactuar con otros servicios. AWS Esta función de servicio debe incluir permisos que permitan a EMR Studio establecer un canal de red seguro entre los espacios de trabajo y los clústeres, almacenar los archivos del bloc de notas y acceder a ellos AWS Secrets Manager al vincular un espacio de trabajo a un repositorio de Git. Amazon S3 Control

Utilice el rol de servicio de Studio (en lugar de las políticas de sesión) para definir todos los permisos de acceso a Amazon S3 para almacenar los archivos de cuadernos y para definir los permisos de acceso de AWS Secrets Manager .

Cómo crear un rol de servicio para EMR Studio en Amazon o EC2 Amazon EKS

1. Siga las instrucciones de Creación de un rol para delegar permisos a un AWS servicio para crear el rol de servicio con la siguiente política de confianza.

   > **Important**
   >
   > La siguiente política de confianza incluye las claves de condición globales `aws:SourceArn` y `aws:SourceAccount` para limitar los permisos que concede a EMR Studio a determinados recursos de su cuenta. Si lo hace, podrá protegerse contra el problema del suplente confuso.

   JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
```

```
        ],
        "Resource": "arn:aws:iam::123456789012:role/EMRStudioServiceRole",
        "Condition": {
          "StringEquals": {
            "aws:SourceAccount": "123456789012"
          },
          "ArnLike": {
            "aws:SourceArn": "arn:aws:elasticmapreduce:*:123456789012:*"
          }
        },
        "Sid": "AllowSTSAssumerole"
      }
    ]
}
```

2. Quite los permisos de rol predeterminados. A continuación, incluya los permisos del siguiente ejemplo de política de permisos de IAM. Como opción, puede crear una política personalizada que usa Permisos del rol de servicio de EMR Studio.

> **Important**
>
> - Para que el control de acceso EC2 basado en etiquetas de Amazon funcione con EMR Studio, debe configurar el acceso a `ModifyNetworkInterfaceAttribute` la API tal y como se muestra en la siguiente política.
> - Para asegurarse de que EMR Studio trabaje con el rol de servicio, no debe cambiar ninguna de estas instrucciones: `AllowAddingEMRTagsDuringDefaultSecurityGroupCreation` y `AllowAddingTagsDuringEC2ENICreation`.
> - Para usar la política de ejemplo, debe etiquetar los siguientes recursos con la clave **`"for-use-with-amazon-emr-managed-policies"`** y el valor **`"true"`**.
> - Su Amazon Virtual Private Cloud (VPC) para EMR Studio.
> - Cada subred que desee utilizar con el estudio.
> - Cualquier grupo de seguridad de EMR Studio personalizado. Debe etiquetar los grupos de seguridad que haya creado durante el período de vista previa de EMR Studio si desea seguir utilizándolos.
> - Secretos guardados en él AWS Secrets Manager que los usuarios de Studio utilizan para vincular los repositorios de Git a un espacio de trabajo.

Puede aplicar etiquetas a los recursos mediante la pestaña Etiquetas de la pantalla de recursos correspondiente de la AWS Management Console.

Cuando proceda, cambie el `*` en `"Resource":"*"` en la siguiente política para especificar el nombre de recurso de Amazon (ARN) de los recursos que la instrucción cubre en sus casos de uso.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowEMRReadOnlyActions",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:ListInstances",
        "elasticmapreduce:DescribeCluster",
        "elasticmapreduce:ListSteps"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "AllowEC2ENIActionsWithEMRTags",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterfacePermission",
        "ec2:DeleteNetworkInterface"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:network-interface/*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
```

```
    {
      "Sid": "AllowEC2ENIAttributeAction",
      "Effect": "Allow",
      "Action": [
        "ec2:ModifyNetworkInterfaceAttribute"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:instance/*",
        "arn:aws:ec2:*:*:network-interface/*",
        "arn:aws:ec2:*:*:security-group/*"
      ]
    },
    {
      "Sid": "AllowEC2SecurityGroupActionsWithEMRTags",
      "Effect": "Allow",
      "Action": [
        "ec2:AuthorizeSecurityGroupEgress",
        "ec2:AuthorizeSecurityGroupIngress",
        "ec2:RevokeSecurityGroupEgress",
        "ec2:RevokeSecurityGroupIngress",
        "ec2:DeleteNetworkInterfacePermission"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "AllowDefaultEC2SecurityGroupsCreationWithEMRTags",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateSecurityGroup"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:security-group/*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
        }
```

```
        }
      },
      {
        "Sid": "AllowDefaultEC2SecurityGroupsCreationInVPCWithEMRTags",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateSecurityGroup"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:vpc/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "AllowAddingEMRTagsDuringDefaultSecurityGroupCreation",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateTags"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:security-group/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true",
            "ec2:CreateAction": "CreateSecurityGroup"
          }
        }
      },
      {
        "Sid": "AllowEC2ENICreationWithEMRTags",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateNetworkInterface"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:network-interface/*"
        ],
        "Condition": {
          "StringEquals": {
```

```
            "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "AllowEC2ENICreationInSubnetAndSecurityGroupWithEMRTags",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateNetworkInterface"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:subnet/*",
          "arn:aws:ec2:*:*:security-group/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "AllowAddingTagsDuringEC2ENICreation",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateTags"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:network-interface/*"
        ],
        "Condition": {
          "StringEquals": {
            "ec2:CreateAction": "CreateNetworkInterface"
          }
        }
      },
      {
        "Sid": "AllowEC2ReadOnlyActions",
        "Effect": "Allow",
        "Action": [
          "ec2:DescribeSecurityGroups",
          "ec2:DescribeNetworkInterfaces",
          "ec2:DescribeTags",
          "ec2:DescribeInstances",
          "ec2:DescribeSubnets",
```

```
        "ec2:DescribeVpcs"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "AllowSecretsManagerReadOnlyActionsWithEMRTags",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetSecretValue"
      ],
      "Resource": [
        "arn:aws:secretsmanager:*:*:secret:*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "AllowWorkspaceCollaboration",
      "Effect": "Allow",
      "Action": [
        "iam:GetUser",
        "iam:GetRole",
        "iam:ListUsers",
        "iam:ListRoles",
        "sso:GetManagedApplicationInstance",
        "sso-directory:SearchUsers"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

3. Conceda a su rol de servicio acceso de lectura y escritura a su ubicación de Amazon S3 para EMR Studio. Utilice el siguiente conjunto mínimo de permisos. Para obtener más información, consulte el ejemplo Amazon S3: permite el acceso de lectura y escritura a objetos en un bucket de S3 mediante programación y en la consola.

```
"s3:PutObject",
"s3:GetObject",
"s3:GetEncryptionConfiguration",
"s3:ListBucket",
"s3:DeleteObject"
```

Si cifra su bucket de Amazon S3, incluya los siguientes permisos para AWS Key Management Service.

```
"kms:Decrypt",
"kms:GenerateDataKey",
"kms:ReEncryptFrom",
"kms:ReEncryptTo",
"kms:DescribeKey"
```

4. Si quiere controlar el acceso a los secretos de Git a nivel de usuario, agregue permisos basados en etiquetas a `secretsmanager:GetSecretValue` en la política de roles de usuario de EMR Studio y elimine los permisos a la política `secretsmanager:GetSecretValue` de la política de roles de servicio de EMR Studio. Para obtener más información acerca de cómo utilizar permisos detallados, consulte Creación de políticas de permisos para los usuarios de EMR Studio.

Rol de servicio mínimo para EMR sin servidor

Si desea ejecutar cargas de trabajo interactivas con EMR sin servidor a través de los cuadernos de EMR Studio, utilice la misma política de confianza que empleó para configurar EMR Studio en la sección anterior: Cómo crear un rol de servicio para EMR Studio en Amazon o EC2 Amazon EKS.

Para su política de IAM, la política mínima viable tiene los siguientes permisos. Actualice *`bucket-name`* con el nombre del bucket que planea usar al configurar su EMR Studio y espacio de trabajo. EMR Studio usa el bucket como copia de seguridad de los espacios de trabajo y los archivos de cuadernos de su estudio.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
```

```
    {
      "Sid": "ObjectActions",
      "Effect": "Allow",
      "Action": [
        "s3:PutObject",
        "s3:GetObject",
        "s3:DeleteObject"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name/*"
      ]
    },
    {
      "Sid": "BucketActions",
      "Effect": "Allow",
      "Action": [
        "s3:ListBucket",
        "s3:GetEncryptionConfiguration"
      ],
      "Resource": [
        "arn:aws:s3:::bucket-name"
      ]
    }
  ]
}
```

Si pretende usar un bucket de Amazon S3 cifrado, agregue los siguientes permisos a su política:

```
"kms:Decrypt",
"kms:GenerateDataKey",
"kms:ReEncryptFrom",
"kms:ReEncryptTo",
"kms:DescribeKey"
```

Permisos del rol de servicio de EMR Studio

En la siguiente tabla se enumeran las operaciones que EMR Studio realiza con el rol de servicio, junto con las acciones de IAM necesarias para cada operación.

Operación	Acciones

Establezca un canal de red seguro entre un espacio de trabajo y un clúster de EMR y lleve a cabo las acciones de limpieza necesarias.	`"ec2:CreateNetworkInterface",`
	`"ec2:CreateNetworkInterfacePermission",`
	`"ec2:DeleteNetworkInterface",`
	`"ec2:DeleteNetworkInterfacePermission",`
	`"ec2:DescribeNetworkInterfaces",`
	`"ec2:ModifyNetworkInterfaceAttribute",`
	`"ec2:AuthorizeSecurityGroupEgress",`
	`"ec2:AuthorizeSecurityGroupIngress",`
	`"ec2:CreateSecurityGroup",`
	`"ec2:DescribeSecurityGroups",`
	`"ec2:RevokeSecurityGroupEgress",`
	`"ec2:DescribeTags",`
	`"ec2:DescribeInstances",`
	`"ec2:DescribeSubnets",`
	`"ec2:DescribeVpcs",`
	`"elasticmapreduce:ListInstances",`
	`"elasticmapreduce:DescribeCluster",`
	`"elasticmapreduce:ListSteps"`

Usa las credenciales de Git almacenadas AWS Secrets Manager para vincular los repositorios de Git a un espacio de trabajo.	`"secretsmanager:GetSecretValue"`
Aplique AWS etiquetas a la interfaz de red y a los grupos de seguridad predeterminados que EMR Studio crea al configurar el canal de red seguro. Para obtener más informaci ón, consulte Tagging AWS resources (Etiquetado de los recursos de).	`"ec2:CreateTags"`

Acceda a los archivos y metadatos de los cuadernos en Amazon S3 o cárguelos.	`"s3:PutObject",`
	`"s3:GetObject",`
	`"s3:GetEncryptionConfiguration",`
	`"s3:ListBucket",`
	`"s3:DeleteObject"`

Operación	Acciones

	Si usa un bucket de Amazon S3 cifrado, incluya los siguientes permisos.
	`"kms:Decrypt",`
	`"kms:GenerateDataKey",`
	`"kms:ReEncryptFrom",`
	`"kms:ReEncryptTo",`
	`"kms:DescribeKey"`

Habilite y configure la colaboración en el espacio de trabajo.	`"iam:GetUser",`
	`"iam:GetRole",`
	`"iam:ListUsers",`
	`"iam:ListRoles",`
	`"sso:GetManagedApplicationInstance",`
	`"sso-directory:SearchUsers",`
	`"sso:DescribeApplication",`
	`"sso:DescribeInstance"`

Cifre las libretas y los archivos del espacio de trabajo de EMR Studio mediante claves administradas por el cliente (CMK) con AWS Key Management Service	`"kms:Decrypt",`
	`"kms:GenerateDataKey",`
	`"kms:ReEncryptFrom",`
	`"kms:ReEncryptTo",`
	`"kms:DescribeKey"`

Configurar los permisos de usuario de EMR Studio para Amazon o EC2 Amazon EKS

Debe configurar las políticas de permisos de usuario para Amazon EMR Studio para poder establecer permisos detallados de usuarios y grupos. Para obtener información sobre cómo funcionan los permisos de usuario en EMR Studio, consulte Control de acceso en Cómo funciona Amazon EMR Studio.

> **Note**
>
> Los permisos que se describen en esta sección no aplican el control de acceso a los datos. Para administrar el acceso a los conjuntos de datos de entrada, debe configurar los permisos

para los clústeres que usa Studio. Para obtener más información, consulte Seguridad en Amazon EMR.

Creación de un rol de usuario de EMR Studio para el modo de autenticación de IAM Identity Center

Debe crear un rol de usuario de EMR Studio cuando usa el modo de autenticación de IAM Identity Center.

Para crear un rol de usuario para EMR Studio

1. Siga las instrucciones de la Guía del AWS Identity and Access Management usuario sobre cómo crear un rol para delegar permisos a un AWS servicio para crear un rol de usuario.

   Al crear el rol, utilice la siguiente política de relaciones de confianza.

   JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole",
        "sts:SetContext"
      ],
      "Resource": "arn:aws:iam::123456789012:role/EMRStudioServiceRole",
      "Sid": "AllowSTSAssumerole"
    }
  ]
}
```

2. Quite los permisos de rol y las políticas predeterminados.
3. Asocie sus políticas de sesión de EMR Studio al rol de usuario antes de asignar usuarios y grupos a un estudio. Para obtener instrucciones sobre cómo crear políticas de sesión, consulte Creación de políticas de permisos para los usuarios de EMR Studio.

Creación de políticas de permisos para los usuarios de EMR Studio

Puede ver las siguientes secciones para crear políticas de permisos para EMR Studio.

Temas

- Creación de las políticas de permisos
- Definir la propiedad de la colaboración en el espacio de trabajo
- Creación de una política de secretos de Git de nivel de usuario
- Asociación de la política de permisos a la identidad de IAM

> **Note**
>
> Para configurar los permisos de acceso a Amazon S3 para almacenar los archivos de cuadernos y establecer los permisos de acceso de AWS Secrets Manager para leer secretos al vincular espacios de trabajo a los repositorios de Git, utilice el rol de servicio de EMR Studio.

Creación de las políticas de permisos

Cree una o varias políticas de permisos de IAM que especifiquen qué acciones puede realizar un usuario en su Studio. Por ejemplo, puede crear tres políticas independientes para los usuarios básicos, intermedios y avanzados de Studio con los ejemplos de políticas de esta página.

Para ver un desglose de todas las operaciones que puede realizar un usuario de Studio junto con las acciones mínimas de IAM necesarias para realizar esa operación, consulte AWS Identity and Access Management permisos para los usuarios de EMR Studio. Para ver los pasos para crear las políticas, consulte Creación de políticas de IAM en la Guía el usuario de IAM.

La política de permisos debe incluir las siguientes instrucciones.

```
{
            "Sid": "AllowAddingTagsOnSecretsWithEMRStudioPrefix",
            "Effect": "Allow",
            "Action": "secretsmanager:TagResource",
            "Resource": "arn:aws:secretsmanager:*:*:secret:emr-studio-*"
},
{
            "Sid": "AllowPassingServiceRoleForWorkspaceCreation",
```

```
            "Action": "iam:PassRole",
            "Resource": [
                "arn:aws:iam::*:role/your-emr-studio-service-role"
            ],
            "Effect": "Allow"
}
```

Definir la propiedad de la colaboración en el espacio de trabajo

La colaboración en el espacio de trabajo permite que varios usuarios trabajen simultáneamente en el mismo espacio de trabajo y se puede configurar con el panel Colaboración de la interfaz de usuario del espacio de trabajo. Para ver y utilizar el panel Colaboración, el usuario debe tener los siguientes permisos. Cualquier usuario con estos permisos puede ver y usar el panel Colaboración.

```
"elasticmapreduce:UpdateEditor",
"elasticmapreduce:PutWorkspaceAccess",
"elasticmapreduce:DeleteWorkspaceAccess",
"elasticmapreduce:ListWorkspaceAccessIdentities"
```

Para restringir el acceso al panel Colaboración, puede utilizar el control de acceso basado en etiquetas. Cuando un usuario crea un espacio de trabajo, EMR Studio aplica una etiqueta predeterminada con una clave de `creatorUserId` cuyo valor es el ID del usuario que crea el espacio de trabajo.

> **Note**
>
> EMR Studio agrega la etiqueta `creatorUserId` a los espacios de trabajo creados después del 16 de noviembre de 2021. Para restringir quién puede configurar la colaboración en espacios de trabajo que usted creó antes, le recomendamos que agregue manualmente la etiqueta `creatorUserId` a su espacio de trabajo y, a continuación, utilice el control de acceso basado en etiquetas en sus políticas de permisos de usuario.

La siguiente instrucción de ejemplo permite a un usuario configurar la colaboración para cualquier espacio de trabajo con la clave de etiqueta `creatorUserId` cuyo valor coincida con el ID del usuario (indicado por la variable de política `aws:userId`). En otras palabras, la instrucción permite al usuario configurar la colaboración para los espacios de trabajo que cree. Para obtener más información sobre variables de políticas, consulte Elementos de la política de IAM: variables y etiquetas en la Guía del usuario de IAM.

```
    {
        "Sid": "UserRolePermissionsForCollaboration",
        "Action": [
            "elasticmapreduce:UpdateEditor",
            "elasticmapreduce:PutWorkspaceAccess",
            "elasticmapreduce:DeleteWorkspaceAccess",
            "elasticmapreduce:ListWorkspaceAccessIdentities"
        ],
        "Resource": "*",
        "Effect": "Allow",
        "Condition": {
            "StringEquals": {
                "elasticmapreduce:ResourceTag/creatorUserId": "${aws:userid}"
            }
        }
    }
```

Creación de una política de secretos de Git de nivel de usuario

Temas

- Para usar permisos de nivel de usuario
- Para pasar de los permisos de nivel de servicio a los permisos de nivel de usuario
- Para usar permisos de nivel de servicio

Para usar permisos de nivel de usuario

EMR Studio agrega automáticamente la etiqueta `for-use-with-amazon-emr-managed-user-policies` cuando crea secretos de Git. Si quiere controlar el acceso a los secretos de Git a nivel de usuario, agregue permisos basados en etiquetas a la política de roles de usuario de EMR Studio con `secretsmanager:GetSecretValue` como se muestra en la siguiente sección Para pasar de los permisos de nivel de servicio a los permisos de nivel de usuario.

Si ya tiene permisos para `secretsmanager:GetSecretValue` en la política de roles de servicio de EMR Studio debe eliminarlos.

Para pasar de los permisos de nivel de servicio a los permisos de nivel de usuario

> **Note**
>
> La etiqueta `for-use-with-amazon-emr-managed-user-policies` garantiza que los permisos del paso 1 a continuación concedan al creador del espacio de trabajo acceso al secreto de Git. Sin embargo, si enlazó los repositorios de Git antes del 1 de septiembre de 2023, se denegará el acceso a los secretos de Git correspondientes porque no tienen la etiqueta `for-use-with-amazon-emr-managed-user-policies` aplicada. Para aplicar permisos a nivel de usuario, debes volver a crear los antiguos secretos JupyterLab y volver a vincular los repositorios de Git correspondientes.
> Para obtener más información sobre las variables de las políticas, consulte Elementos de la política de IAM: variables y etiquetas en la Guía del usuario de IAM.

1. Agregue los siguientes permisos a la política de roles de usuario de EMR Studio. Utiliza la clave `for-use-with-amazon-emr-managed-user-policies` con el valor `"${aws:userid}"`.

```
{
   "Sid": "AllowSecretsManagerReadOnlyActionsWithEMRTags",
    "Effect": "Allow",
    "Action": "secretsmanager:GetSecretValue",
    "Resource": "arn:aws:secretsmanager:*:*:secret:*",
    "Condition": {
        "StringEquals": {
            "secretsmanager:ResourceTag/for-use-with-amazon-emr-managed-user-
policies": "${aws:userid}"
        }
    }
}
```

2. Si está presente, elimine el siguiente permiso de la política de roles de servicio de EMR Studio. Como la política de rol de servicio se aplica a todos los secretos definidos por cada usuario, solo tiene que hacerlo una vez.

```
{
    "Sid": "AllowSecretsManagerReadOnlyActionsWithEMRTags",
    "Effect": "Allow",
    "Action": [
        "secretsmanager:GetSecretValue"
```

```
        ],
        "Resource": "arn:aws:secretsmanager:*:*:secret:*",
        "Condition": {
            "StringEquals": {
                "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
            }
        }
    }
```

Para usar permisos de nivel de servicio

A partir del 1 de septiembre de 2023, EMR Studio agrega automáticamente la etiqueta `for-use-with-amazon-emr-managed-user-policies` para el control de acceso a nivel de usuario. Como se trata de una capacidad adicional, puede seguir utilizando el acceso a nivel de servicio que está disponible mediante el permiso `GetSecretValue` en el rol de servicio de EMR Studio.

Para los secretos creados antes del 1 de septiembre de 2023, EMR Studio no agregó la etiqueta `for-use-with-amazon-emr-managed-user-policies`. Para seguir utilizando los permisos de nivel de servicio, simplemente conserve sus permisos actuales de rol de servicio de EMR Studio y rol de usuario. Sin embargo, para restringir quién puede acceder a un secreto individual, le recomendamos que siga los pasos en Para usar permisos de nivel de usuario para agregar manualmente la etiqueta `for-use-with-amazon-emr-managed-user-policies` a sus secretos y, a continuación, utilice el control de acceso basado en etiquetas en sus políticas de permisos de usuario.

Para obtener más información sobre las variables de las políticas, consulte Elementos de la política de IAM: variables y etiquetas en la Guía del usuario de IAM.

Asociación de la política de permisos a la identidad de IAM

En la siguiente tabla se resume la identidad de IAM a la que se asocia una política de permisos, según el modo de autenticación de EMR Studio. Para obtener instrucciones acerca de cómo asociar una política, consulte Agregar y quitar permisos de identidad de IAM.

Si usa…	Asociar la política a…
Autenticación de IAM	Sus identidades de IAM (usuarios, grupos de usuarios o roles). Por ejemplo, puede asociar

Si usa…	Asociar la política a…
	una política de permisos a un usuario en su Cuenta de AWS.
Federación de IAM con un proveedor de identidades (IdP) externo	El rol o roles de IAM que cree para su IdP externo. Por ejemplo, una federación de IAM para SAML 2.0. EMR Studio usa los permisos que asocia a sus roles de IAM para los usuarios con acceso federado a un estudio.
IAM Identity Center	Su rol de usuario de Amazon EMR Studio.

Ejemplo de políticas de usuario

La siguiente política de usuario básico permite la mayoría de las acciones de EMR Studio, pero no permite que un usuario cree nuevos clústeres de Amazon EMR.

Política básica

> **Important**
>
> La política de ejemplo no incluye el permiso `CreateStudioPresignedUrl`, que debe conceder a un usuario al utilizar el modo de autenticación de IAM. Para obtener más información, consulte Asignar un usuario o grupo a un EMR Studio.

La política de ejemplo incluye elementos `Condition` para aplicar el control de acceso basado en etiquetas (TBAC) de modo que pueda utilizar la política con el rol de servicio de ejemplo para EMR Studio. Para obtener más información, consulte Crear un rol de servicio de EMR Studio.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
```

```
        "Sid": "AllowDefaultEC2SecurityGroupsCreationInVPCWithEMRTags",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateSecurityGroup"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:vpc/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "AllowAddingEMRTagsDuringDefaultSecurityGroupCreation",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateTags"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:security-group/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true",
            "ec2:CreateAction": "CreateSecurityGroup"
          }
        }
      },
      {
        "Sid": "AllowSecretManagerListSecrets",
        "Action": [
          "secretsmanager:ListSecrets"
        ],
        "Resource": [
          "*"
        ],
        "Effect": "Allow"
      },
      {
        "Sid": "AllowSecretCreationWithEMRTagsAndEMRStudioPrefix",
        "Effect": "Allow",
        "Action": [
```

```
        "secretsmanager:CreateSecret"
      ],
      "Resource": [
        "arn:aws:secretsmanager:*:*:secret:emr-studio-*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "AllowAddingTagsOnSecretsWithEMRStudioPrefix",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:TagResource"
      ],
      "Resource": [
        "arn:aws:secretsmanager:*:*:secret:emr-studio-*"
      ]
    },
    {
      "Sid": "AllowPassingServiceRoleForWorkspaceCreation",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/your-emr-studio-service-role>"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowS3ListAndLocationPermissions",
      "Action": [
        "s3:ListAllMyBuckets",
        "s3:ListBucket",
        "s3:GetBucketLocation"
      ],
      "Resource": [
        "arn:aws:s3:::*"
      ],
      "Effect": "Allow"
    },
    {
```

```
        "Sid": "AllowS3ReadOnlyAccessToLogs",
        "Action": [
          "s3:GetObject"
        ],
        "Resource": [
          "arn:aws:s3:::aws-logs-aws-111122223333>-region>/elasticmapreduce/*"
        ],
        "Effect": "Allow"
      },
      {
        "Sid": "AllowConfigurationForWorkspaceCollaboration",
        "Action": [
          "elasticmapreduce:UpdateEditor",
          "elasticmapreduce:PutWorkspaceAccess",
          "elasticmapreduce:DeleteWorkspaceAccess",
          "elasticmapreduce:ListWorkspaceAccessIdentities"
        ],
        "Resource": [
          "*"
        ],
        "Effect": "Allow",
        "Condition": {
          "StringEquals": {
            "elasticmapreduce:ResourceTag/creatorUserId": "${aws:userId}"
          }
        }
      },
      {
        "Sid": "DescribeNetwork",
        "Effect": "Allow",
        "Action": [
          "ec2:DescribeVpcs",
          "ec2:DescribeSubnets",
          "ec2:DescribeSecurityGroups"
        ],
        "Resource": [
          "*"
        ]
      },
      {
        "Sid": "ListIAMRoles",
        "Effect": "Allow",
        "Action": [
          "iam:ListRoles"
```

```
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

La siguiente política de usuario intermedio permite la mayoría de las acciones de EMR Studio y permite al usuario crear nuevos clústeres de Amazon EMR mediante una plantilla de clústeres.

Política intermedia

> **Important**
>
> La política de ejemplo no incluye el permiso `CreateStudioPresignedUrl`, que debe conceder a un usuario al utilizar el modo de autenticación de IAM. Para obtener más información, consulte Asignar un usuario o grupo a un EMR Studio.

La política de ejemplo incluye elementos `Condition` para aplicar el control de acceso basado en etiquetas (TBAC) de modo que pueda utilizar la política con el rol de servicio de ejemplo para EMR Studio. Para obtener más información, consulte Crear un rol de servicio de EMR Studio.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowEMRBasicActions",
      "Action": [
        "elasticmapreduce:CreateEditor",
        "elasticmapreduce:DescribeEditor",
        "elasticmapreduce:ListEditors",
        "elasticmapreduce:StartEditor",
        "elasticmapreduce:StopEditor",
        "elasticmapreduce:DeleteEditor",
        "elasticmapreduce:OpenEditorInConsole",
        "elasticmapreduce:AttachEditor",
```

```
        "elasticmapreduce:DetachEditor",
        "elasticmapreduce:CreateRepository",
        "elasticmapreduce:DescribeRepository",
        "elasticmapreduce:DeleteRepository",
        "elasticmapreduce:ListRepositories",
        "elasticmapreduce:LinkRepository",
        "elasticmapreduce:UnlinkRepository",
        "elasticmapreduce:DescribeCluster",
        "elasticmapreduce:ListInstanceGroups",
        "elasticmapreduce:ListBootstrapActions",
        "elasticmapreduce:ListClusters",
        "elasticmapreduce:ListSteps",
        "elasticmapreduce:CreatePersistentAppUI",
        "elasticmapreduce:DescribePersistentAppUI",
        "elasticmapreduce:GetPersistentAppUIPresignedURL",
        "elasticmapreduce:GetOnClusterAppUIPresignedURL"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowEMRContainersBasicActions",
      "Action": [
        "emr-containers:DescribeVirtualCluster",
        "emr-containers:ListVirtualClusters",
        "emr-containers:DescribeManagedEndpoint",
        "emr-containers:ListManagedEndpoints",
        "emr-containers:DescribeJobRun",
        "emr-containers:ListJobRuns"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowRetrievingManagedEndpointCredentials",
      "Effect": "Allow",
      "Action": [
        "emr-containers:GetManagedEndpointSessionCredentials"
      ],
      "Resource": [
```

```
        "arn:aws:emr-containers:us-west-1:123456789012:/virtualclusters/virtual-
cluster-id/endpoints/managed-endpoint-id"
      ],
      "Condition": {
        "StringEquals": {
          "emr-containers:ExecutionRoleArn": [
            "arn:aws:iam::123456789012:role/emr-on-eks-execution-role"
          ]
        }
      }
    },
    {
      "Sid": "AllowSecretManagerListSecrets",
      "Action": [
        "secretsmanager:ListSecrets"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowSecretCreationWithEMRTagsAndEMRStudioPrefix",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:CreateSecret"
      ],
      "Resource": [
        "arn:aws:secretsmanager:*:*:secret:emr-studio-*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "AllowAddingTagsOnSecretsWithEMRStudioPrefix",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:TagResource"
      ],
      "Resource": [
        "arn:aws:secretsmanager:*:*:secret:emr-studio-*"
```

```
        ]
      },
      {
        "Sid": "AllowClusterTemplateRelatedIntermediateActions",
        "Action": [
          "servicecatalog:DescribeProduct",
          "servicecatalog:DescribeProductView",
          "servicecatalog:DescribeProvisioningParameters",
          "servicecatalog:ProvisionProduct",
          "servicecatalog:SearchProducts",
          "servicecatalog:UpdateProvisionedProduct",
          "servicecatalog:ListProvisioningArtifacts",
          "servicecatalog:ListLaunchPaths",
          "servicecatalog:DescribeRecord",
          "cloudformation:DescribeStackResources"
        ],
        "Resource": [
          "*"
        ],
        "Effect": "Allow"
      },
      {
        "Sid": "AllowPassingServiceRoleForWorkspaceCreation",
        "Action": [
          "iam:PassRole"
        ],
        "Resource": [
          "arn:aws:iam::*:role/your-emr-studio-service-role"
        ],
        "Effect": "Allow"
      },
      {
        "Sid": "AllowS3ListAndLocationPermissions",
        "Action": [
          "s3:ListAllMyBuckets",
          "s3:ListBucket",
          "s3:GetBucketLocation"
        ],
        "Resource": [
          "arn:aws:s3:::*"
        ],
        "Effect": "Allow"
      },
      {
```

```
        "Sid": "AllowS3ReadOnlyAccessToLogs",
        "Action": [
          "s3:GetObject"
        ],
        "Resource": [
          "arn:aws:s3:::aws-logs-123456789012-us-east-1/elasticmapreduce/*"
        ],
        "Effect": "Allow"
      },
      {
        "Sid": "AllowConfigurationForWorkspaceCollaboration",
        "Action": [
          "elasticmapreduce:UpdateEditor",
          "elasticmapreduce:PutWorkspaceAccess",
          "elasticmapreduce:DeleteWorkspaceAccess",
          "elasticmapreduce:ListWorkspaceAccessIdentities"
        ],
        "Resource": [
          "*"
        ],
        "Effect": "Allow",
        "Condition": {
          "StringEquals": {
            "elasticmapreduce:ResourceTag/creatorUserId": "${aws:userId}"
          }
        }
      },
      {
        "Sid": "DescribeNetwork",
        "Effect": "Allow",
        "Action": [
          "ec2:DescribeVpcs",
          "ec2:DescribeSubnets",
          "ec2:DescribeSecurityGroups"
        ],
        "Resource": [
          "*"
        ]
      },
      {
        "Sid": "ListIAMRoles",
        "Effect": "Allow",
        "Action": [
          "iam:ListRoles"
```

```
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "AllowServerlessActions",
      "Action": [
        "emr-serverless:CreateApplication",
        "emr-serverless:UpdateApplication",
        "emr-serverless:DeleteApplication",
        "emr-serverless:ListApplications",
        "emr-serverless:GetApplication",
        "emr-serverless:StartApplication",
        "emr-serverless:StopApplication",
        "emr-serverless:StartJobRun",
        "emr-serverless:CancelJobRun",
        "emr-serverless:ListJobRuns",
        "emr-serverless:GetJobRun",
        "emr-serverless:GetDashboardForJobRun",
        "emr-serverless:AccessInteractiveEndpoints"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowPassingRuntimeRoleForRunningServerlessJob",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/serverless-runtime-role"
      ],
      "Effect": "Allow"
    }
  ]
}
```

La siguiente política de usuario avanzada admite todas las acciones de EMR Studio y permite al usuario crear nuevos clústeres de Amazon EMR mediante una plantilla de clústeres o proporcionando una configuración de clúster.

Política avanzada

> **Important**
>
> La política de ejemplo no incluye el permiso `CreateStudioPresignedUrl`, que debe conceder a un usuario al utilizar el modo de autenticación de IAM. Para obtener más información, consulte Asignar un usuario o grupo a un EMR Studio.

La política de ejemplo incluye elementos `Condition` para aplicar el control de acceso basado en etiquetas (TBAC) de modo que pueda utilizar la política con el rol de servicio de ejemplo para EMR Studio. Para obtener más información, consulte Crear un rol de servicio de EMR Studio.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowEMRBasicActions",
      "Action": [
        "elasticmapreduce:CreateEditor",
        "elasticmapreduce:DescribeEditor",
        "elasticmapreduce:ListEditors",
        "elasticmapreduce:StartEditor",
        "elasticmapreduce:StopEditor",
        "elasticmapreduce:DeleteEditor",
        "elasticmapreduce:OpenEditorInConsole",
        "elasticmapreduce:AttachEditor",
        "elasticmapreduce:DetachEditor",
        "elasticmapreduce:CreateRepository",
        "elasticmapreduce:DescribeRepository",
        "elasticmapreduce:DeleteRepository",
        "elasticmapreduce:ListRepositories",
        "elasticmapreduce:LinkRepository",
        "elasticmapreduce:UnlinkRepository",
        "elasticmapreduce:DescribeCluster",
```

```
        "elasticmapreduce:ListInstanceGroups",
        "elasticmapreduce:ListBootstrapActions",
        "elasticmapreduce:ListClusters",
        "elasticmapreduce:ListSteps",
        "elasticmapreduce:CreatePersistentAppUI",
        "elasticmapreduce:DescribePersistentAppUI",
        "elasticmapreduce:GetPersistentAppUIPresignedURL",
        "elasticmapreduce:GetOnClusterAppUIPresignedURL"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowEMRContainersBasicActions",
      "Action": [
        "emr-containers:DescribeVirtualCluster",
        "emr-containers:ListVirtualClusters",
        "emr-containers:DescribeManagedEndpoint",
        "emr-containers:ListManagedEndpoints",
        "emr-containers:DescribeJobRun",
        "emr-containers:ListJobRuns"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowRetrievingManagedEndpointCredentials",
      "Effect": "Allow",
      "Action": [
        "emr-containers:GetManagedEndpointSessionCredentials"
      ],
      "Resource": [
        "arn:aws:emr-containers:*:123456789012:/virtualclusters/virtual-cluster-
id/endpoints/managed-endpoint-id"
      ],
      "Condition": {
        "StringEquals": {
          "emr-containers:ExecutionRoleArn": [
            "arn:aws:iam::123456789012:role/emr-on-eks-execution-role"
          ]
```

```
        }
      }
    },
    {
      "Sid": "AllowSecretManagerListSecrets",
      "Action": [
        "secretsmanager:ListSecrets"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowSecretCreationWithEMRTagsAndEMRStudioPrefix",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:CreateSecret"
      ],
      "Resource": [
        "arn:aws:secretsmanager:*:*:secret:emr-studio-*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "AllowAddingTagsOnSecretsWithEMRStudioPrefix",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:TagResource"
      ],
      "Resource": [
        "arn:aws:secretsmanager:*:*:secret:emr-studio-*"
      ]
    },
    {
      "Sid": "AllowClusterTemplateRelatedIntermediateActions",
      "Action": [
        "servicecatalog:DescribeProduct",
        "servicecatalog:DescribeProductView",
        "servicecatalog:DescribeProvisioningParameters",
```

```
        "servicecatalog:ProvisionProduct",
        "servicecatalog:SearchProducts",
        "servicecatalog:UpdateProvisionedProduct",
        "servicecatalog:ListProvisioningArtifacts",
        "servicecatalog:ListLaunchPaths",
        "servicecatalog:DescribeRecord",
        "cloudformation:DescribeStackResources"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowEMRCreateClusterAdvancedActions",
      "Action": [
        "elasticmapreduce:RunJobFlow"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowPassingServiceRoleForWorkspaceCreation",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/your-emr-studio-service-role",
        "arn:aws:iam::*:role/EMR_DefaultRole_V2",
        "arn:aws:iam::*:role/EMR_EC2_DefaultRole"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowS3ListAndLocationPermissions",
      "Action": [
        "s3:ListAllMyBuckets",
        "s3:ListBucket",
        "s3:GetBucketLocation"
      ],
      "Resource": [
        "arn:aws:s3:::*"
```

```
        ],
        "Effect": "Allow"
      },
      {
        "Sid": "AllowS3ReadOnlyAccessToLogs",
        "Action": [
          "s3:GetObject"
        ],
        "Resource": [
          "arn:aws:s3:::aws-logs-123456789012-us-east-1/elasticmapreduce/*"
        ],
        "Effect": "Allow"
      },
      {
        "Sid": "AllowConfigurationForWorkspaceCollaboration",
        "Action": [
          "elasticmapreduce:UpdateEditor",
          "elasticmapreduce:PutWorkspaceAccess",
          "elasticmapreduce:DeleteWorkspaceAccess",
          "elasticmapreduce:ListWorkspaceAccessIdentities"
        ],
        "Resource": [
          "*"
        ],
        "Effect": "Allow",
        "Condition": {
          "StringEquals": {
            "elasticmapreduce:ResourceTag/creatorUserId": "${aws:userId}"
          }
        }
      },
      {
        "Sid": "SageMakerDataWranglerForEMRStudio",
        "Effect": "Allow",
        "Action": [
          "sagemaker:CreatePresignedDomainUrl",
          "sagemaker:DescribeDomain",
          "sagemaker:ListDomains",
          "sagemaker:ListUserProfiles"
        ],
        "Resource": [
          "*"
        ]
      },
```

```
    {
      "Sid": "DescribeNetwork",
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeVpcs",
        "ec2:DescribeSubnets",
        "ec2:DescribeSecurityGroups"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "ListIAMRoles",
      "Effect": "Allow",
      "Action": [
        "iam:ListRoles"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "AllowServerlessActions",
      "Action": [
        "emr-serverless:CreateApplication",
        "emr-serverless:UpdateApplication",
        "emr-serverless:DeleteApplication",
        "emr-serverless:ListApplications",
        "emr-serverless:GetApplication",
        "emr-serverless:StartApplication",
        "emr-serverless:StopApplication",
        "emr-serverless:StartJobRun",
        "emr-serverless:CancelJobRun",
        "emr-serverless:ListJobRuns",
        "emr-serverless:GetJobRun",
        "emr-serverless:GetDashboardForJobRun",
        "emr-serverless:AccessInteractiveEndpoints"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
```

```
    {
      "Sid": "AllowPassingRuntimeRoleForRunningServerlessJob",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/serverless-runtime-role"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowCodeWhisperer",
      "Effect": "Allow",
      "Action": [
        "codewhisperer:GenerateRecommendations"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "AllowAthenaSQL",
      "Action": [
        "athena:StartQueryExecution",
        "athena:StopQueryExecution",
        "athena:GetQueryExecution",
        "athena:GetQueryRuntimeStatistics",
        "athena:GetQueryResults",
        "athena:ListQueryExecutions",
        "athena:BatchGetQueryExecution",
        "athena:GetNamedQuery",
        "athena:ListNamedQueries",
        "athena:BatchGetNamedQuery",
        "athena:UpdateNamedQuery",
        "athena:DeleteNamedQuery",
        "athena:ListDataCatalogs",
        "athena:GetDataCatalog",
        "athena:ListDatabases",
        "athena:GetDatabase",
        "athena:ListTableMetadata",
        "athena:GetTableMetadata",
        "athena:ListWorkGroups",
        "athena:GetWorkGroup",
        "athena:CreateNamedQuery",
```

```
        "athena:GetPreparedStatement",
        "glue:CreateDatabase",
        "glue:DeleteDatabase",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:UpdateDatabase",
        "glue:CreateTable",
        "glue:DeleteTable",
        "glue:BatchDeleteTable",
        "glue:UpdateTable",
        "glue:GetTable",
        "glue:GetTables",
        "glue:BatchCreatePartition",
        "glue:CreatePartition",
        "glue:DeletePartition",
        "glue:BatchDeletePartition",
        "glue:UpdatePartition",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition",
        "kms:ListAliases",
        "kms:ListKeys",
        "kms:DescribeKey",
        "lakeformation:GetDataAccess",
        "s3:GetObject",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:ListMultipartUploadParts",
        "s3:AbortMultipartUpload",
        "s3:PutObject",
        "s3:PutBucketPublicAccessBlock",
        "s3:ListAllMyBuckets"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    }
  ]
}
```

La siguiente política de usuario contiene los permisos de usuario básicos que se necesitan para utilizar una aplicación interactiva de EMR sin servidor con espacios de trabajo de EMR Studio.

Política interactiva de EMR sin servidor

En este ejemplo de política que tiene permisos de usuario para las aplicaciones interactivas EMR Serverless con EMR Studio, sustituya los marcadores de posición por y por el rol de servicio EMR Studio y el *serverless-runtime-role* rol de tiempo de *emr-studio-service-role* ejecución EMR Serverless correctos.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowServerlessActions",
      "Action": [
        "emr-serverless:CreateApplication",
        "emr-serverless:UpdateApplication",
        "emr-serverless:DeleteApplication",
        "emr-serverless:ListApplications",
        "emr-serverless:GetApplication",
        "emr-serverless:StartApplication",
        "emr-serverless:StopApplication",
        "emr-serverless:StartJobRun",
        "emr-serverless:CancelJobRun",
        "emr-serverless:ListJobRuns",
        "emr-serverless:GetJobRun",
        "emr-serverless:GetDashboardForJobRun",
        "emr-serverless:AccessInteractiveEndpoints"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowEMRBasicActions",
      "Action": [
        "elasticmapreduce:CreateEditor",
        "elasticmapreduce:DescribeEditor",
```

```
        "elasticmapreduce:ListEditors",
        "elasticmapreduce:UpdateStudio",
        "elasticmapreduce:StartEditor",
        "elasticmapreduce:StopEditor",
        "elasticmapreduce:DeleteEditor",
        "elasticmapreduce:OpenEditorInConsole",
        "elasticmapreduce:AttachEditor",
        "elasticmapreduce:DetachEditor",
        "elasticmapreduce:CreateStudio",
        "elasticmapreduce:DescribeStudio",
        "elasticmapreduce:DeleteStudio",
        "elasticmapreduce:ListStudios",
        "elasticmapreduce:CreateStudioPresignedUrl"
      ],
      "Resource": [
        "*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowPassingRuntimeRoleForRunningEMRServerlessJob",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/serverless-runtime-role"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowPassingServiceRoleForWorkspaceCreation",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/emr-studio-service-role"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "AllowS3ListAndGetPermissions",
      "Action": [
        "s3:ListAllMyBuckets",
        "s3:ListBucket",
```

```
        "s3:GetBucketLocation",
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::*"
      ],
      "Effect": "Allow"
    },
    {
      "Sid": "DescribeNetwork",
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeVpcs",
        "ec2:DescribeSubnets",
        "ec2:DescribeSecurityGroups"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "ListIAMRoles",
      "Effect": "Allow",
      "Action": [
        "iam:ListRoles"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

AWS Identity and Access Management permisos para los usuarios de EMR Studio

En la tabla siguiente se indican todas las operaciones de Amazon EMR Studio que puede realizar un usuario y se enumeran las acciones de IAM mínimas necesarias para realizar esa operación. Puede permitir estas acciones en sus políticas de permisos de IAM (al usar la autenticación de IAM) o en sus políticas de sesión con roles de usuario (al usar la autenticación de IAM Identity Center) para EMR Studio.

En la tabla también se muestran las operaciones permitidas en cada una de las políticas de permisos de ejemplo para EMR Studio. Para obtener más información acerca de las políticas de permisos de ejemplo, consulte Creación de políticas de permisos para los usuarios de EMR Studio.

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Crear y eliminar espacios de trabajo	Sí	Sí	Sí	`"elasticmapreduce:CreateEditor", "elasticmapreduce:DescribeEditor", "elasticmapreduce:ListEditors", "elasticmapreduce:DeleteEditor"`
Consulte el panel Colaboración, habilite la colaboración en el espacio de trabajo y agregue a colaboradores. Para obtener más información, consulte Establecer la propiedad de la colaboración en el espacio de trabajo.	Sí	Sí	Sí	`"elasticmapreduce:UpdateEditor", "elasticmapreduce:PutWorkspaceAccess", "elasticmapreduce:DeleteWorkspaceAccess", "elasticmapreduce:ListWorkspaceAccessIdentities"`
Consulte una lista de depósitos de Amazon S3 Control almacenamiento en la misma cuenta que Studio al crear un nuevo clúster de EMR y acceda a los registros del contenedor cuando utilice una interfaz de usuario	Sí	Sí	Sí	`"s3:ListAllMyBuckets", "s3:ListBucket", "s3:GetBucketLocation", "s3:GetObject"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
web para depurar aplicaciones				
Acceder a los espacios de trabajo	Sí	Sí	Sí	`"elasticmapreduce:DescribeEditor", "elasticmapreduce:ListEditors", "elasticmapreduce:StartEditor", "elasticmapreduce:StopEditor", "elasticmapreduce:OpenEditorInConsole"`
Asociar o separar los clústeres de Amazon EMR existentes asociados al espacio de trabajo	Sí	Sí	Sí	`"elasticmapreduce:AttachEditor", "elasticmapreduce:DetachEditor", "elasticmapreduce:ListClusters", "elasticmapreduce:DescribeCluster", "elasticmapreduce:ListInstanceGroups", "elasticmapreduce:ListBootstrapActions"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Asociar o separar clústeres de Amazon EMR en EKS	Sí	Sí	Sí	`"elasticmapreduce:AttachEditor", "elasticmapreduce:DetachEditor", "emr-containers:ListVirtualClusters", "emr-containers:DescribeVirtualCluster", "emr-containers:ListManagedEndpoints", "emr-containers:DescribeManagedEndpoint", "emr-containers:GetManagedEndpointSessionCredentials"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Adjuntar o desconectar las aplicaciones de EMR sin servidor asociadas al espacio de trabajo	No	Sí	Sí	`"elasticmapreduce:AttachEditor", "elasticmapreduce:DetachEditor", "emr-serverless:GetApplication", "emr-serverless:StartApplication", "emr-serverless:ListApplications", "emr-serverless:GetDashboardForJobRun", "emr-serverless:AccessInteractiveEndpoints", "iam:PassRole"` El permiso `PassRole` es necesario para pasar el rol de tiempo de ejecución del trabajo de EMR sin servidor. Para obtener más información, consulte Roles de tiempo de ejecución de trabajo en la Guía del usuario de Amazon EMR sin servidor.

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Depure Amazon EMR EC2 en trabajos con interfaces de usuario de aplicaciones persistentes	Sí	Sí	Sí	`"elasticmapreduce:CreatePersistentAppUI", "elasticmapreduce:DescribePersistentAppUI", "elasticmapreduce:GetPersistentAppUIPresignedURL", "elasticmapreduce:ListClusters", "elasticmapreduce:ListSteps", "elasticmapreduce:DescribeCluster", "s3:ListBucket", "s3:GetObject"`
Depure Amazon EMR EC2 en trabajos con interfaces de usuario de aplicaciones en el clúster	Sí	Sí	Sí	`"elasticmapreduce:GetOnClusterAppUIPresignedURL"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Depurar Amazon EMR en ejecuciones de trabajos de EKS mediante el servidor de historial de Spark	Sí	Sí	Sí	`"elasticmapreduce:CreatePersistentAppUI", "elasticmapreduce:DescribePersistentAppUI", "elasticmapreduce:GetPersistentAppUIPresignedURL", "emr-containers:ListVirtualClusters", "emr-containers:DescribeVirtualCluster", "emr-containers:ListJobRuns", "emr-containers:DescribeJobRun", "s3:ListBucket", "s3:GetObject"`
Crear y eliminar repositorios de Git	Sí	Sí	Sí	`"elasticmapreduce:CreateRepository", "elasticmapreduce:DeleteRepository", "elasticmapreduce:ListRepositories", "elasticmapreduce:DescribeRepository", "secretsmanager:CreateSecret", "secretsmanager:ListSecrets", "secretsmanager:TagResource"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Vincular y desvincular repositorios de Git	Sí	Sí	Sí	`"elasticmapreduce:LinkRepository", "elasticmapreduce:UnlinkRepository", "elasticmapreduce:ListRepositories", "elasticmapreduce:DescribeRepository"`
Crear nuevos clústeres a partir de plantillas de clústeres predefinidas	No	Sí	Sí	`"servicecatalog:SearchProducts", "servicecatalog:DescribeProduct", "servicecatalog:DescribeProductView", "servicecatalog:DescribeProvisioningParameters", "servicecatalog:ProvisionProduct", "servicecatalog:UpdateProvisionedProduct", "servicecatalog:ListProvisioningArtifacts", "servicecatalog:DescribeRecord", "servicecatalog:ListLaunchPaths", "cloudformation:DescribeStackResources", "elasticmapreduce:ListClusters", "elasticmapreduce:DescribeCluster"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Brinde una configuración de clúster para crear nuevos clústeres.	No	No	Sí	`"elasticmapreduce:RunJobFlow", "iam:PassRole", "elasticmapreduce:ListClusters", "elasticmapreduce:DescribeCluster"`
Asigne un usuario a un Studio cuando utilice el modo de autenticación de IAM.	No	No	No	`"elasticmapreduce:CreateStudioPresignedUrl"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Describir los objetos de la red.	Sí	Sí	Sí	JSON `{ "Version": "2012-10-17", "Statement": [ { "Sid": "DescribeNetwork", "Effect": "Allow", "Action": [ "ec2:DescribeVpcs", "ec2:DescribeSubnets", "ec2:DescribeSecurityGroups" ], "Resource": [ "*" ] } ] }`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Enumerar los roles de IAM.	Sí	Sí	Sí	JSON `{ "Version": "2012-10-17", "Statement": [ { "Sid": "ListIAMRoles", "Effect": "Allow", "Action": [ "iam:ListRoles" ], "Resource": [ "*" ] } ] }`
Conéctese a EMR Studio desde Amazon SageMaker AI Studio y utilice la interfaz visual Data Wrangler.	No	No	Sí	`"sagemaker:CreatePresignedDomainUrl", "sagemaker:DescribeDomain", "sagemaker:ListDomains", "sagemaker:ListUserProfiles"`
Utilice Amazon CodeWhisperer en su estudio de EMR.	No	No	Sí	`"codewhisperer:GenerateRecommendations"`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
Acceda al editor SQL de Amazon Athena desde EMR Studio. Es posible que esta lista no incluya todos los permisos que necesita para usar todas las características de Athena. Para obtener la mayor parte up-to-date de la lista, consulte la política de acceso completo de Athena.	No	No	Sí	`"athena:StartQueryExecution", "athena:StopQueryExecution", "athena:GetQueryExecution", "athena:GetQueryRuntimeStatistics", "athena:GetQueryResults", "athena:ListQueryExecutions", "athena:BatchGetQueryExecution", "athena:GetNamedQuery", "athena:ListNamedQueries", "athena:BatchGetNamedQuery", "athena:UpdateNamedQuery", "athena:DeleteNamedQuery", "athena:ListDataCatalogs", "athena:GetDataCatalog", "athena:ListDatabases", "athena:GetDatabase", "athena:ListTableMetadata", "athena:GetTableMetadata", "athena:ListWorkGroups", "athena:GetWorkGroup", "athena:CreateNamedQuery", "athena:GetPreparedStatement", "glue:CreateDatabase", "glue:DeleteDatabase", "glue:GetDatabase",`

Acción	Basic	Intermedia	Avanzado	Acciones asociadas
				"glue:GetDatabases", "glue:UpdateDatabase", "glue:CreateTable", "glue:DeleteTable", "glue:BatchDeleteTable", "glue:UpdateTable", "glue:GetTable", "glue:GetTables", "glue:BatchCreatePartition", "glue:CreatePartition", "glue:DeletePartition", "glue:BatchDeletePartition", "glue:UpdatePartition", "glue:GetPartition", "glue:GetPartitions", "glue:BatchGetPartition", "kms:ListAliases", "kms:ListKeys", "kms:DescribeKey", "lakeformation:GetDataAccess", "s3:GetBucketLocation", "s3:GetBucketLocation", "s3:GetObject", "s3:ListBucket", "s3:ListBucketMultipartUploads", "s3:ListMultipartUploadParts", "s3:AbortMultipartUpload", "s3:PutObject", "s3:PutBucketPublicAccessBlock", "s3:ListAllMyBuckets"

Crear un EMR Studio

Puede crear un EMR Studio para su equipo con la consola de Amazon EMR o la AWS CLI. La creación de una instancia de Studio forma parte de la configuración de Amazon EMR Studio.

Requisitos previos

Antes de crear un estudio, asegúrese de haber completado las tareas anteriores de Configurar un EMR Studio.

Para crear un estudio con AWS CLI, debe tener instalada la última versión. Para obtener más información, consulte Instalación o actualización de la versión de AWS CLI más reciente.

> **Important**
>
> Desactive las herramientas de administración de proxy, como FoxyProxy o SwitchyOmega en el navegador, antes de crear un Studio. Los proxies activos pueden provocar un mensaje de error de red al seleccionar Crear estudio.

Amazon EMR le proporciona una experiencia de consola sencilla para crear un Studio, de forma que pueda empezar rápidamente con la configuración predeterminada para ejecutar cargas de trabajo interactivas o trabajos por lotes con la configuración predeterminada. La creación de un EMR Studio también crea una aplicación EMR sin servidor lista para sus trabajos interactivos.

Si desea tener el control total de los ajustes de su Studio, puede elegir Personalizado, lo que le permitirá configurar todos los ajustes adicionales.

Interactive workloads

Para crear un EMR Studio para cargas de trabajo interactivas

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR Studio, en el panel de navegación de la izquierda, elija Comenzar. También puede crear un estudio nuevo desde la página Studios.
3. Amazon EMR le proporciona la configuración predeterminada si va a crear un EMR Studio para cargas de trabajo interactivas, pero puede editar esta configuración. Los ajustes configurables incluyen el nombre de EMR Studio, la ubicación S3 de su espacio de trabajo, el rol de servicio que debe usar, los espacios de trabajo que desea usar, el nombre de la aplicación EMR sin servidor y el rol de tiempo de ejecución asociado.

4. Seleccione Crear un Studio y lanzar un espacio de trabajo para terminar y vaya a la página Studios. El nuevo estudio aparecerá en la lista con detalles como el nombre del estudio, la fecha de creación y la URL de acceso al estudio. Se abrirá su espacio de trabajo en una pestaña nueva en su navegador.

Batch jobs

Para crear un EMR Studio para cargas de trabajo interactivas

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR Studio, en el panel de navegación de la izquierda, elija Comenzar. También puede crear un estudio nuevo desde la página Studios.
3. Amazon EMR le proporciona la configuración predeterminada si va a crear un EMR Studio para trabajos por lotes, pero puede editar esta configuración. Los ajustes configurables incluyen el nombre de EMR Studio, el nombre de la aplicación EMR sin servidor y el rol de tiempo de ejecución asociado.
4. Seleccione Crear un Studio y lanzar un espacio de trabajo para terminar y vaya a la página Studios. El nuevo estudio aparecerá en la lista con detalles como el nombre del estudio, la fecha de creación y la URL de acceso al estudio. Se abrirá su EMR Studio en una pestaña nueva en su navegador.

Custom settings

Para crear un EMR Studio con ajustes personalizados

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR Studio, en el panel de navegación de la izquierda, elija Comenzar. También puede crear un estudio nuevo desde la página Studios.
3. Seleccione Crear un estudio para abrir la página Crear un estudio.
4. Introduzca un nombre de Studio.
5. Elija crear un bucket de S3 nuevo o utilizar una ubicación existente.
6. Elija el espacio de trabajo que desee añadir al Studio. Puede agregar hasta 3 espacios de trabajo.

7. En Autenticación, seleccione un modo de autenticación para el estudio y proporcione la información según la siguiente tabla. Para obtener más información sobre la autenticación en EMR Studio, consulte Elija un modo de autenticación para Amazon EMR Studio.

Si usa…	Haga lo siguiente...
Autenticación o federación de IAM	El método de autenticación predeterminado es AWS Identity and Access Managemen t (IAM). En la parte inferior de la pantalla, también puede agregar etiquetas para que los usuarios específicos puedan acceder al Studio, como se describe en Asignar un usuario o grupo a un EMR Studio. Si desea que los usuarios federados inicien sesión con la URL de Studio y las credenciales de su proveedor de identidad (IdP), seleccione su IdP en la lista desplegable e introduzca la URL de inicio de sesión del proveedor de identidad (IdP) y el nombre del parámetro. RelayState Para obtener una lista de RelayState nombres URLs y autenticaciones de IdP, consulte. RelayState Parámetros y autenticación del proveedor de identidades URLs

Si usa…	Haga lo siguiente...
Autenticación del Centro de identidades de IAM	Seleccione su rol de servicio y rol de usuario de EMR Studio. Para obtener más información, consulte Crear un rol de servicio de EMR Studio y Creación de un rol de usuario de EMR Studio para el modo de autenticación de IAM Identity Center. Si utiliza la autenticación del IAM Identity Center (anteriormente AWS Single Sign On) para Studio, puede optar por agilizar la experiencia de inicio de sesión de los usuarios con la opción Habilitar la propagación de identidades de forma fiable. Gracias a la propagación de identidad es fiable, los usuarios pueden iniciar sesión con sus credenciales del Centro de Identidad y hacer que sus identidades se propaguen a los AWS servicios descenden tes cuando utilizan Studio. En la sección Acceso a las aplicacio nes, también puede especificar si todos los usuarios y grupos de Identity Center pueden tener acceso al Studio o si solo los usuarios y grupos asignados que elija pueden acceder al Studio. Para obtener más informaciónIntegre Amazon EMR con AWS IAM Identity Center, consulte la Guía del usuario del AWS IAM Identity Center y también sobre la propagación fiable de la identidad entre aplicaciones.

8. En VPC, seleccione una Amazon Virtual Private Cloud (VPC) para el Studio en la lista desplegable.

9. En Subredes, seleccione un máximo de cinco subredes en tu VPC para asociarlas al estudio. Tiene la opción de agregar más subredes después de crear el estudio.
10. En Grupos de seguridad, seleccione los grupos de seguridad predeterminados o grupos de seguridad personalizados. Para obtener más información, consulte Definir grupos de seguridad para controlar el tráfico de red de EMR Studio.

Si selecciona...	Haga lo siguiente...
Los grupos de seguridad predeterminados de EMR Studio	Para habilitar la vinculación de repositor ios basada en Git para Studio, seleccion a Habilitar clusters/endpoints y repositorio de Git. De lo contrario, seleccione Habilitar clústeres/puntos de conexión.
Grupos de seguridad personalizados para su estudio	• En Grupo de seguridad de los clústeres/ puntos de conexión, seleccione el grupo de seguridad del motor que configuró en la lista desplegable. Su estudio usa este grupo de seguridad para permitir el acceso entrante desde los espacios de trabajo asociados. • En Grupo de seguridad del espacio de trabajo, seleccione el grupo de seguridad del espacio de trabajo que configuró en la lista desplegable. Su estudio usa este grupo de seguridad con los espacios de trabajo para proporcionar acceso saliente a los clústeres de Amazon EMR asociados y a los repositorios de Git alojados públicamente.

11. Añadir etiquetas a su Studio y otros recursos. Para obtener más información acerca de las etiquetas, consulte Clústeres de equitetas.
12. Seleccione Crear un Studio y lanzar un espacio de trabajo para terminar y vaya a la página Studios. El nuevo estudio aparecerá en la lista con detalles como el nombre del estudio, la fecha de creación y la URL de acceso al estudio.

Después de crear el estudio, siga las instrucciones que aparecen en Asignar un usuario o grupo a un EMR Studio.

CLI

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Example - Creación de un EMR Studio que utilice IAM para la autenticación

El siguiente AWS CLI comando de ejemplo crea un EMR Studio con el modo de autenticación de IAM. Cuando utiliza la autenticación o federación de IAM del estudio, no tiene que especificar un `--user-role`.

Para permitir que los usuarios federados inicien sesión con la URL de Studio y las credenciales de su proveedor de identidades (IdP), especifique su `--idp-auth-url` y `--idp-relay-state-parameter-name`. Para obtener una lista de RelayState nombres URLs y autenticaciones de IdP, consulte. RelayState Parámetros y autenticación del proveedor de identidades URLs

```
aws emr create-studio \
--name <example-studio-name> \
--auth-mode IAM \
--vpc-id <example-vpc-id> \
--subnet-ids <subnet-id-1> <subnet-id-2>... <subnet-id-5>  \
--service-role <example-studio-service-role-name> \
--user-role studio-user-role-name \
--workspace-security-group-id <example-workspace-sg-id> \
--engine-security-group-id <example-engine-sg-id> \
--default-s3-location <example-s3-location> \
--idp-auth-url <https://EXAMPLE/login/> \
--idp-relay-state-parameter-name <example-RelayState>
```

Example - Creación de un EMR Studio que utilice Identity Center para la autenticación

El siguiente comando de AWS CLI ejemplo crea un EMR Studio que utiliza el modo de autenticación de IAM Identity Center. Al utilizar la autenticación de IAM Identity Center, debe especificar un `--user-role`.

Para obtener más información acerca del modo de autenticación de IAM Identity Center, consulte Configuración del modo de autenticación de IAM Identity Center para Amazon EMR Studio.

```
aws emr create-studio \
--name <example-studio-name> \
--auth-mode SSO \
--vpc-id <example-vpc-id> \
--subnet-ids <subnet-id-1> <subnet-id-2>... <subnet-id-5>  \
--service-role <example-studio-service-role-name> \
--user-role <example-studio-user-role-name> \
--workspace-security-group-id <example-workspace-sg-id> \
--engine-security-group-id <example-engine-sg-id> \
--default-s3-location <example-s3-location>
--trusted-identity-propagation-enabled \
--idc-user-assignment OPTIONAL \
--idc-instance-arn <iam-identity-center-instance-arn>
```

Example - Salida de la CLI para **`aws emr create-studio`**

A continuación, se muestra un ejemplo de la salida que aparece después de crear un estudio.

```
{
    StudioId: "es-123XXXXXXXXX",
    Url: "https://es-123XXXXXXXXX.emrstudio-prod.us-east-1.amazonaws.com"
}
```

Para obtener más información sobre el comando `create-studio`, consulte Referencia del comando AWS CLI .

RelayState Parámetros y autenticación del proveedor de identidades URLs

Si utilizas la federación de IAM y quieres que los usuarios inicien sesión con la URL de Studio y las credenciales de tu proveedor de identidad (IdP), puedes especificar la URL de inicio de sesión de tu proveedor de identidad (IdP) RelayStatey el nombre del parámetro cuando lo hagas. Crear un EMR Studio

En la siguiente tabla se muestran la URL de autenticación estándar y el nombre del RelayState parámetro de algunos proveedores de identidad populares.

Proveedor de identidades	Parámetro	URL de autenticación
Auth0	`RelayState`	`https://<sub_domain> .auth0.com/samlp/<app_id>`
Cuentas de Google	`RelayState`	`https://accounts.google.com/o/saml2/initsso?idpid= <idp_id>&spid=<sp_id>&forceauthn=false`
Microsoft Azure	`RelayState`	`https://myapps.microsoft.com/signin/ <app_name> /<app_id>?tenantId= <tenant_id>`
Okta	`RelayState`	`https://<sub_domain> .okta.com/app/<app_name> /<app_id>/sso/saml`
PingFederate	`TargetResource`	`https://<host>/idp/<idp_id>/startSSO.ping?PartnerSpId= <sp_id>`
PingOne	`TargetResource`	`https://sso.connect.pingidentity.com/sso/sp/initsso?saasid= <app_id>&idpid=<idp_id>`

Asignar y administrar usuarios de EMR Studio

Después de crear un EMR Studio, puede asignarle usuarios y grupos. El método que utilice para asignar, actualizar y eliminar usuarios depende del modo de autenticación de Studio.

- Cuando utiliza el modo de autenticación de IAM, debe configurar la asignación de usuarios y los permisos de EMR Studio en IAM o con IAM y su proveedor de identidades.
- Con el modo de autenticación de IAM Identity Center, utiliza la consola de administración de Amazon EMR o AWS CLI la para gestionar los usuarios.

Para obtener más información sobre la autenticación en Amazon EMR Studio, consulte Elija un modo de autenticación para Amazon EMR Studio.

Asignar un usuario o grupo a un EMR Studio

IAM

Cuando utilice Configuración del modo de autenticación de IAM para Amazon EMR Studio, debe permitir la acción `CreateStudioPresignedUrl` en la política de permisos de IAM de un usuario y restringirlo a un estudio concreto. Puede incluir `CreateStudioPresignedUrl` en su Permisos de usuario para el modo de autenticación de IAM o utilizar una política independiente.

Para restringir a un usuario a un estudio (o conjunto de estudios), puede usar el control de acceso basado en atributos (ABAC) o especificar el nombre de recurso de Amazon (ARN) de un estudio en el elemento `Resource` de la política de permisos.

Example Asignar un usuario a un estudio mediante el ARN del estudio

El siguiente ejemplo de política proporciona a un usuario acceso a un EMR Studio concreto al permitir la acción `CreateStudioPresignedUrl` y especificar el nombre de recurso de Amazon (ARN) del estudio en el elemento `Resource`.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowCreateStudioPresignedUrl",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:CreateStudioPresignedUrl"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:us-east-1:123456789012:studio/studio-id"
      ]
    }
  ]
}
```

Example Asignar un usuario a un estudio con ABAC para la autenticación de IAM

Hay varias formas de configurar el control de acceso basado en atributos (ABAC) en un estudio. Por ejemplo, puede asociar una o más etiquetas a un EMR Studio y, a continuación, crear una política de IAM que restrinja la acción `CreateStudioPresignedUrl` a un estudio concreto o a un conjunto de estudios con esas etiquetas.

Puede agregar etiquetas durante o después de la creación del estudio. Para agregar etiquetas a un estudio existente, puede utilizar el comando AWS CLI`emr add-tags`. El siguiente ejemplo agrega una etiqueta con el par clave-valor `Team = Data Analytics` a un EMR Studio.

```
aws emr add-tags --resource-id <example-studio-id> --tags Team="Data Analytics"
```

El siguiente ejemplo de política de permisos permite la acción `CreateStudioPresignedUrl` para los estudios de EMR Studio con el par clave-valor de etiqueta `Team = DataAnalytics`. Para obtener más información sobre el uso de etiquetas para el control de acceso, consulte Controlar el acceso a y para los usuarios y roles utilizando etiquetas o Controlar el acceso a los recursos de AWS utilizando etiquetas.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowCreateStudioPresignedUrl",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:CreateStudioPresignedUrl"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:studio/*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/Team": "Data Analytics"
        }
      }
    }
  ]
```

```
}
```

Example Asigne un usuario a un estudio mediante la clave de condición aws: SourceIdentity global

Al utilizar la federación de IAM, puede utilizar la clave de condición global `aws:SourceIdentity` en una política de permisos para que los usuarios accedan a Studio cuando asuman su rol de IAM para la federación.

Primero debe configurar su proveedor de identidades (IdP) para que devuelva una cadena de identificación, como una dirección de correo electrónico o un nombre de usuario, cuando un usuario se autentique y asuma su rol de IAM para la federación. IAM establece la clave de condición global `aws:SourceIdentity` en la cadena de identificación devuelta por el IdP.

Para obtener más información, consulte la entrada del blog Cómo relacionar la actividad del rol de IAM con la identidad corporativa en el blog de AWS seguridad y la referencia aws: SourceIdentity entry in the global condition keys.

El siguiente ejemplo de política permite la `CreateStudioPresignedUrl` acción y proporciona a los usuarios un acceso `aws:SourceIdentity` que coincide con el *`<example-source-identity>`* acceso al EMR Studio especificado por. *`<example-studio-arn>`*

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:CreateStudioPresignedUrl"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:us-east-1:123456789012:studio/studio-name"
      ],
      "Condition": {
        "StringLike": {
          "aws:SourceIdentity": "example-source-identity"
        }
      },
```

```
        "Sid": "AllowELASTICMAPREDUCECreatestudiopresignedurl"
      }
    ]
}
```

IAM Identity Center

Al asignar un usuario o un grupo a un EMR Studio, debe especificar una política de sesión que define permisos detallados, como la capacidad de crear un nuevo clúster de EMR, para ese usuario o grupo. Amazon EMR almacena estas asignaciones de políticas de sesión. Puede actualizar la política de sesión de un usuario o grupo después de la asignación.

> **Note**
>
> Los permisos finales de un usuario o grupo son una intersección de los permisos definidos en su rol de usuario de EMR Studio y los permisos definidos en la política de sesión para ese usuario o grupo. Si un usuario pertenece a más de un grupo asignado al estudio, EMR Studio utiliza una unión de permisos para ese usuario.

Para asignar usuarios o grupos a un EMR Studio mediante la consola de Amazon EMR

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Seleccione EMR Studio en el menú de navegación de la izquierda.
3. Elija el nombre de su estudio en la lista Studios o seleccione el estudio y elija Ver detalles para abrir la página de detalles del estudio.
4. Elija la pestaña Agregar usuarios para ver la tabla de búsqueda Usuarios y Grupos.
5. Seleccione la pestaña Usuarios o la pestaña Grupos e introduzca un término de búsqueda en la barra de búsqueda para buscar un usuario o grupo.
6. Seleccione uno o más usuarios o grupos de la lista de resultados de la búsqueda. Puede cambiar de una pestaña a otra de Usuarios y Grupos.
7. Tras seleccionar los usuarios y grupos para agregarlos al estudio, seleccione Agregar. Debería ver los usuarios y grupos en la lista Usuarios del estudio. La lista puede tardar unos segundos en actualizarse.

8. Siga las instrucciones que se indican en Actualizar los permisos de un usuario o grupo asignado a un estudio para ajustar los permisos de Studio de un usuario o un grupo.

Para asignar un usuario o grupo a un EMR Studio con la AWS CLI

Inserte sus propios valores para los siguientes argumentos de `create-studio-session-mapping`. Para obtener más información sobre el comando `create-studio-session-mapping`, consulte la Referencia de comandos de la AWS CLI .

- **`--studio-id`**: el ID del estudio al que quiere asignar el usuario o grupo. Para obtener instrucciones sobre cómo recuperar un ID de Studio, consulte Ver detalles del estudio.
- **`--identity-name`**: el nombre del usuario o grupo del Almacén de identidades. Para obtener más información, consulte la UserNamesección sobre usuarios y DisplayNamegrupos en la referencia de la API de Identity Store.
- **`--identity-type`**: utilice `USER` o `GROUP` para especificar el tipo de identidad.
- **`--session-policy-arn`**: el Nombre de recurso de Amazon (ARN) de la política de sesión que desee asociar al usuario o grupo. Por ejemplo, **`arn:aws:iam::`*`<aws-account-id>`*`:policy/`*`EMRStudio_Advanced_User_Policy`**. Para obtener más información, consulte Creación de políticas de permisos para los usuarios de EMR Studio.

```
aws emr create-studio-session-mapping \
 --studio-id <example-studio-id> \
 --identity-name <example-identity-name> \
 --identity-type <USER-or-GROUP> \
 --session-policy-arn <example-session-policy-arn>
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

Utilice el comando `get-studio-session-mapping` para comprobar la nueva asignación. *`<example-identity-name>`*Sustitúyalo por el nombre del centro de identidad de IAM del usuario o grupo que actualizaste.

```
aws emr get-studio-session-mapping \
 --studio-id <example-studio-id> \
 --identity-type <USER-or-GROUP> \
 --identity-name <user-or-group-name> \
```

Actualizar los permisos de un usuario o grupo asignado a un estudio

IAM

Para actualizar los permisos de un usuario o grupo al utilizar el modo de autenticación de IAM, debe utilizar IAM para cambiar las políticas de permisos de IAM asociadas a sus identidades de IAM (usuarios, grupos o roles).

Para obtener más información, consulte Permisos de usuario para el modo de autenticación de IAM.

IAM Identity Center

Para actualizar los permisos de EMR Studio de un usuario o grupo mediante la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Seleccione EMR Studio en el menú de navegación de la izquierda.
3. Elija el nombre de su estudio en la lista Studios o seleccione el estudio y elija Ver detalles para abrir la página de detalles del estudio.
4. En la lista Usuarios del estudio de la página de detalles del estudio, busque el usuario o grupo que desee actualizar. Puede buscar por nombre o tipo de identidad.
5. Seleccione el usuario o grupo que desee actualizar y seleccione Asignar política para abrir el cuadro de diálogo Política de sesión.
6. Seleccione una política para aplicarla al usuario o grupo que eligió en el paso 5 y seleccione Aplicar política. La lista Usuarios del estudio mostrará el nombre de la política en la columna Política de sesión del usuario o grupo que haya actualizado.

Para actualizar los permisos de EMR Studio para un usuario o grupo mediante el AWS CLI

Inserte sus propios valores para los siguientes argumentos de `update-studio-session-mappings`. Para obtener más información sobre el comando `update-studio-session-mappings`, consulte la Referencia de comandos de la AWS CLI .

```
aws emr update-studio-session-mapping \
 --studio-id <example-studio-id> \
 --identity-name <name-of-user-or-group-to-update> \
 --session-policy-arn <new-session-policy-arn-to-apply> \
 --identity-type <USER-or-GROUP> \
```

Utilice el comando `get-studio-session-mapping` para comprobar la nueva asignación de la política de sesión. *`<example-identity-name>`*Sustitúyalo por el nombre del centro de identidad de IAM del usuario o grupo que actualizó.

```
aws emr get-studio-session-mapping \
 --studio-id <example-studio-id> \
 --identity-type <USER-or-GROUP> \
 --identity-name <user-or-group-name> \
```

Eliminar un usuario o grupo de un estudio

IAM

Para eliminar un usuario o un grupo de un EMR Studio al utilizar el modo de autenticación de IAM, debe revocar el acceso del usuario al estudio reconfigurando la política de permisos de IAM del usuario.

En el siguiente ejemplo de política, suponga que tiene un EMR Studio con el par clave-valor de etiqueta `Team = Quality Assurance`. Según la política, el usuario puede acceder a los estudios etiquetados con la clave `Team` cuyo valor sea igual a `Data Analytics` o `Quality Assurance`. Para eliminar al usuario del estudio etiquetado con `Team = Quality Assurance`, elimine `Quality Assurance` de la lista de valores de etiqueta.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
```

```
    {
      "Sid": "AllowCreateStudioPresignedUrl",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:CreateStudioPresignedUrl"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:us-east-1:123456789012:studio/*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/Team": [
            "Data Analytics",
            "Quality Assurance"
          ]
        }
      }
    }
  ]
}
```

IAM Identity Center

Para eliminar un usuario o un grupo de un EMR Studio mediante la consola

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Seleccione EMR Studio en el menú de navegación de la izquierda.
3. Elija el nombre de su estudio en la lista Studios o seleccione el estudio y elija Ver detalles para abrir la página de detalles del estudio.
4. En la lista Usuarios del estudio, en la página de detalles del estudio, busque el usuario o grupo que desee eliminar del estudio. Puede buscar por nombre o tipo de identidad.
5. Seleccione el usuario o grupo que desea eliminar, seleccione Eliminar y confirme la eliminación. El usuario o grupo que ha eliminado desaparece de la lista Usuarios del estudio.

Para eliminar un usuario o un grupo de un EMR Studio mediante la AWS CLI

Inserte sus propios valores para los siguientes argumentos de `delete-studio-session-mapping`. Para obtener más información sobre el comando `delete-studio-session-mapping`, consulte la Referencia de comandos de la AWS CLI .

```
aws emr delete-studio-session-mapping \
 --studio-id <example-studio-id> \
 --identity-type <USER-or-GROUP> \
 --identity-name <name-of-user-or-group-to-delete> \
```

Supervise, actualice y elimine los recursos de Amazon EMR Studio

En esta sección se incluyen instrucciones que le ayudarán a monitorear, actualizar o eliminar un recurso de EMR Studio. Para obtener información sobre la asignación de usuarios o la actualización de los permisos de los usuarios, consulte Asignar y administrar usuarios de EMR Studio.

Ver detalles del estudio

Console

Para ver detalles sobre un EMR Studio con la nueva consola

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En la barra de navegación de la izquierda, en EMR Studio, seleccione Studios.
3. Seleccione el estudio en la lista Studios para abrir la página de detalles del estudio. En la página de detalles del estudio se incluye información sobre los ajustes del estudio, como la descripción, la VPC y las subredes del estudio.

CLI

Para recuperar los detalles de un estudio EMR por ID de estudio mediante el AWS CLI

Utilice el siguiente `describe-studio` AWS CLI comando para obtener información detallada sobre un estudio de EMR en particular. Para obtener más información, consulte Referencia de comandos de la AWS CLI .

```
aws emr describe-studio \
 --studio-id <id-of-studio-to-describe> \
```

Para recuperar una lista de los estudios de EMR Studio mediante la AWS CLI

Use el siguiente comando `list-studios` AWS CLI . Para obtener más información, consulte Referencia de comandos de la AWS CLI .

```
aws emr list-studios
```

A continuación, se muestra un ejemplo del valor devuelto para el comando `list-studios` en formato JSON.

```
{
    "Studios": [
        {
            "AuthMode": "IAM",
            "VpcId": "vpc-b21XXXXX",
            "Name": "example-studio-name",
            "Url": "https://es-7HWP74SNGDXXXXXXXXXXXXXX.emrstudio-prod.us-
east-1.amazonaws.com",
            "CreationTime": 1605672582.781,
            "StudioId": "es-7HWP74SNGDXXXXXXXXXXXXXX",
            "Description": "example studio description"
        }
    ]
}
```

Monitorear las acciones de Amazon EMR Studio

Ver la actividad de EMR Studio y la API

EMR Studio está integrado con AWS CloudTrail un servicio que proporciona un registro de las acciones realizadas por un usuario, un rol de IAM u otro AWS servicio de EMR Studio. CloudTrail captura las llamadas a la API de EMR Studio como eventos. Puede ver los eventos mediante la CloudTrail consola en https://console.aws.amazon.com/cloudtrail/.

Los eventos de EMR Studio proporcionan información como qué estudio o usuario de IAM realiza una solicitud y qué tipo de solicitud es.

> **Note**
>
> Las acciones realizadas en el clúster, como la ejecución de tareas de cuaderno, no emiten AWS CloudTrail.

También puede crear un registro para la entrega continua de CloudTrail eventos de EMR Studio a un bucket de Amazon S3. Para obtener más información, consulte la Guía del usuario de AWS CloudTrail.

Ejemplo de CloudTrail evento: un usuario llama a la API DescribeStudio

El siguiente es un ejemplo de AWS CloudTrail evento que se crea cuando un usuario llama a la DescribeStudioAPI. `admin` CloudTrail registra el nombre de usuario como`admin`.

> **Note**
>
> Para proteger los detalles de Studio, el evento de la API de EMR Studio para DescribeStudio excluye un valor para. `responseElements`

```
{
   "eventVersion":"1.08",
   "userIdentity":{
      "type":"IAMUser",
      "principalId":"AIDXXXXXXXXXXXXXXXXX",
      "arn":"arn:aws:iam::653XXXXXXXXX:user/admin",
      "accountId":"653XXXXXXXXX",
      "accessKeyId":"AKIAIOSFODNN7EXAMPLE",
      "userName":"admin"
   },
   "eventTime":"2021-01-07T19:13:58Z",
   "eventSource":"elasticmapreduce.amazonaws.com",
   "eventName":"DescribeStudio",
   "awsRegion":"us-east-1",
   "sourceIPAddress":"72.XX.XXX.XX",
   "userAgent":"aws-cli/1.18.188 Python/3.8.5 Darwin/18.7.0 botocore/1.19.28",
   "requestParameters":{
      "studioId":"es-9O5XXXXXXXXXXXXXXXXXXXXX"
   },
   "responseElements":null,
```

```
    "requestID":"0fxxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx",
    "eventID":"b0xxxxxx-xxxx-xxxx-xxxx-xxxxxxxxxxxx",
    "readOnly":true,
    "eventType":"AwsApiCall",
    "managementEvent":true,
    "eventCategory":"Management",
    "recipientAccountId":"653XXXXXXXXX"
}
```

Consulta de la actividad de los trabajos y los usuarios de Spark

Para ver la actividad de trabajos de Spark realizada por los usuarios de Amazon EMR Studio, puede configurar la suplantación de usuarios en un clúster. Con la suplantación de usuarios, cada trabajo de Spark que se envía desde un espacio de trabajo se asocia al usuario de Studio que ejecutó el código.

Cuando la suplantación de usuarios está habilitada, Amazon EMR crea un directorio de usuarios de HDFS en el nodo principal del clúster para cada usuario que ejecuta código en el espacio de trabajo. Por ejemplo, si el usuario `studio-user-1@example.com` ejecuta código, puede conectarse al nodo principal y comprobar si `hadoop fs -ls /user` tiene un directorio para `studio-user-1@example.com`.

Para configurar la suplantación de usuarios de Spark, establezca las siguientes propiedades en las siguientes clasificaciones de configuración:

- `core-site`
- `livy-conf`

```
[
    {
        "Classification": "core-site",
        "Properties": {
          "hadoop.proxyuser.livy.groups": "*",
          "hadoop.proxyuser.livy.hosts": "*"
        }
    },
    {
        "Classification": "livy-conf",
        "Properties": {
          "livy.impersonation.enabled": "true"
```

```
        }
    }
]
```

Para consultar las páginas del servidor de historial, consulte Depurar aplicaciones y trabajos con EMR Studio. También puede conectarse al nodo principal del clúster mediante SSH para ver las interfaces web de la aplicación. Para obtener más información, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

Actualizar un Amazon EMR Studio

Tras crear un EMR Studio, puede actualizar los siguientes atributos mediante la AWS CLI:

- Nombre
- Descripción
- Ubicación de S3 predeterminada
- Subredes

Para actualizar un EMR Studio mediante el AWS CLI

Utilice el `update-studio` AWS CLI comando para actualizar un estudio de EMR. Para obtener más información, consulte Referencia de comandos de la AWS CLI .

> **Note**
>
> Un estudio se puede utilizar con un máximo de 5 subredes. Estas subredes deben pertenecer a la misma VPC que el estudio. La lista de subredes IDs que envíe al `update-studio` comando puede incluir una nueva subred IDs, pero también debe incluir toda la subred IDs que ya haya asociado al Studio. No se pueden eliminar subredes de un estudio.

```
aws emr update-studio \
 --studio-id <example-studio-id-to-update> \
 --name <example-new-studio-name> \
 --subnet-ids <old-subnet-id-1 old-subnet-id-2 old-subnet-id-3 new-subnet-id> \
```

Para verificar los cambios, usa el `describe-studio` AWS CLI comando y especifica tu ID de Studio. Para obtener más información, consulte Referencia de comandos de la AWS CLI .

```
aws emr describe-studio \
 --studio-id <id-of-updated-studio> \
```

Eliminar un Amazon EMR Studio y un espacio de trabajo

Al eliminar un estudio, EMR Studio elimina todas las asignaciones de usuarios y grupos de IAM Identity Center asociadas al estudio.

> **Note**
>
> Al eliminar un estudio, Amazon EMR no elimina los espacios de trabajo asociados a ese estudio. Debe eliminar los espacios de trabajo de su estudio por separado.

Eliminar espacios de trabajo

Console

Dado que cada espacio de trabajo de EMR Studio es una instancia de cuaderno de EMR, puede utilizar la consola de administración de Amazon EMR para eliminar los espacios de trabajo. Puede eliminar espacios de trabajo mediante la consola de Amazon EMR antes o después de eliminar su estudio.

Para eliminar un espacio de trabajo mediante la consola de Amazon EMR

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Seleccione Cuadernos.
3. Seleccione los espacios de trabajo que desea eliminar.
4. Seleccione Eliminar y, a continuación, Eliminar de nuevo para confirmar la acción.
5. Siga las instrucciones para eliminar objetos en la Guía del usuario de la consola de Amazon Simple Storage Service para eliminar los archivos de cuadernos asociados al espacio de trabajo eliminado de Amazon S3.

EMR Studio UI

From the Workspace UI

Eliminar un espacio de trabajo y sus archivos de copia de seguridad asociados de EMR Studio

1. Inicie sesión en su EMR Studio con la URL de acceso de su estudio y elija Espacios de trabajo en el menú de navegación de la izquierda.
2. Busque su espacio de trabajo en la lista y seleccione la casilla de verificación que aparece junto a su nombre. Puede seleccionar varios espacios de trabajo para eliminarlos al mismo tiempo.
3. Seleccione Eliminar en la esquina superior derecha de la lista de Espacios de trabajo y confirme que desea eliminar los espacios de trabajo seleccionados. Elija Eliminar para confirmar.
4. Si desea eliminar los archivos de cuadernos que estaban asociados al espacio de trabajo eliminado de Amazon S3, siga las instrucciones para eliminar objetos en la Guía del usuario de la consola de Amazon Simple Storage Service. Si usted no creó el estudio, consulte a su administrador del estudio para determinar la ubicación de la copia de seguridad de Amazon S3 para el espacio de trabajo eliminado.

From the Workspaces list

Eliminar un espacio de trabajo y sus archivos de copia de seguridad asociados de la lista de espacios de trabajo

1. Navegue hasta la lista Espacios de trabajo en la consola.
2. Seleccione el espacio de trabajo que desee eliminar de la lista y, a continuación, seleccione Acciones.
3. Elija Eliminar.
4. Si desea eliminar los archivos de cuadernos que estaban asociados al espacio de trabajo eliminado de Amazon S3, siga las instrucciones para eliminar objetos en la Guía del usuario de la consola de Amazon Simple Storage Service. Si usted no creó el estudio, consulte a su administrador del estudio para determinar la ubicación de la copia de seguridad de Amazon S3 para el espacio de trabajo eliminado.

Eliminar un EMR Studio

Console

Para eliminar un EMR Studio con la nueva consola

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En la barra de navegación de la izquierda, en EMR Studio, seleccione Studios.
3. Seleccione el estudio en la lista Studios con el botón situado a la izquierda del nombre del estudio. Elija Eliminar.

Old console

Para eliminar un EMR Studio con la antigua consola

1. Abre la consola Amazon EMR en https://console.aws.amazon.com/elasticmapreduce/ casa.
2. Seleccione EMR Studio en el menú de navegación de la izquierda.
3. Seleccione el estudio en la lista Studios y elija Eliminar.

CLI

Para eliminar un estudio de EMR con el AWS CLI

Utilice el `delete-studio` AWS CLI comando para eliminar un estudio de EMR. Para obtener más información, consulte Referencia de comandos de la AWS CLI .

```
aws emr delete-studio --studio-id <id-of-studio-to-delete>
```

Cifrado de cuadernos y archivos del espacio de trabajo de EMR Studio

En un EMR Studio, puede crear y configurar diferentes espacios de trabajo para organizar y ejecutar cuadernos. Estos espacios de trabajo almacenan cuadernos y archivos relacionados en el bucket de Amazon S3 que especifique. De forma predeterminada, estos archivos se cifran con claves administradas por Amazon S3 (SSE-S3) con el cifrado del lado del servidor como el nivel básico de cifrado. También puede optar por utilizar claves KMS administradas por el cliente (SSE-KMS) para cifrar sus archivos. Puede hacerlo mediante la consola de administración de Amazon EMR o mediante el AWS SDK AWS CLI and al crear un EMR Studio.

El cifrado del almacenamiento del espacio de trabajo de EMR Studio está disponible en todas las Regiones en las que está disponible EMR Studio.

Requisitos previos

Antes de poder cifrar el bloc de notas y los archivos del espacio de trabajo de EMR Studio, debe crear AWS Key Management Service una clave de administrador de clientes (CMK) simétrica en la Cuenta de AWS misma región y región que su EMR Studio.

Su política de recursos AWS KMS debe tener los permisos de acceso necesarios para su función de servicio de EMR Studio. El siguiente es un ejemplo de política de IAM que otorga permisos de acceso mínimos para el cifrado de almacenamiento de un espacio de trabajo de EMR Studio:

```
{
    "Sid": "AllowEMRStudioServiceRoleAccess",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::<ACCOUNT_ID>:role/<ROLE_NAME>"
    },
    "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey",
        "kms:ReEncryptFrom",
        "kms:ReEncryptTo",
        "kms:DescribeKey"
    ],
    "Resource": "*",
    "Condition": {
        "StringEquals": {
            "kms:CallerAccount": "<ACCOUNT_ID>",
            "kms:EncryptionContext:aws:s3:arn": "arn:aws:s3:::<S3_BUCKET_NAME>",
            "kms:ViaService": "s3.<AWS_REGION>.amazonaws.com"
        }
    }
}
```

Su rol de servicio de EMR Studio también debe tener los permisos de acceso para usar su AWS KMS clave. El siguiente es un ejemplo de política de IAM que concede los permisos de acceso mínimos para el cifrado de almacenamiento de un espacio de trabajo de EMR Studio:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowEMRStudioWorkspaceStorageEncryptionAccess",
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey",
        "kms:ReEncryptFrom",
        "kms:ReEncryptTo",
        "kms:DescribeKey"
      ],
      "Resource": [
        "arn:aws:kms:*:123456789012:key/12345678-1234-1234-1234-123456789012"
      ]
    }
  ]
}
```

Crear un nuevo EMR Studio

Siga estos pasos para crear un nuevo EMR Studio que utilice el cifrado de almacenamiento del espacio de trabajo.

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com/elasticmapreduce/.
2. Seleccione Studios y, a continuación, seleccione Crear Studio.
3. En el caso de la ubicación de S3 para el almacenamiento, ingrese o elija una ruta de Amazon S3. Esta es la ubicación en Amazon S3 en la que Amazon EMR almacena los cuadernos y archivos del espacio de trabajo.
4. En Rol de servicio elija un rol de IAM. Este es el rol de IAM que asume Amazon EMR.
5. Elija Cifrar los archivos del espacio de trabajo con su propia clave. AWS KMS
6. Introduzca o elija una AWS KMS clave para cifrar las libretas y los archivos del espacio de trabajo en Amazon S3.
7. Elija Crear Studio o Crear Studio y Lanzar espacios de trabajo.

8. Elija Cifrar los archivos del espacio de trabajo con su propia clave. AWS KMS
9. Introduzca o elija una de AWS KMS las que desee utilizar para cifrar las libretas y los archivos del espacio de trabajo en Amazon S3.
10. Elija Save changes (Guardar cambios).

Los siguientes pasos muestran cómo actualizar un EMR Studio y configurar el cifrado del almacenamiento del espacio de trabajo.

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com/elasticmapreduce/.
2. Elija un EMR Studio existente de la lista y, a continuación, seleccione Editar.
3. Elija Cifrar los archivos del espacio de trabajo con su propia clave. AWS KMS
4. Introduzca o elija una de AWS KMS las que desee utilizar para cifrar las libretas y los archivos del espacio de trabajo en Amazon S3.
5. Elija Save changes (Guardar cambios).

Definir grupos de seguridad para controlar el tráfico de red de EMR Studio

Acerca de los grupos de seguridad de EMR Studio

Amazon EMR Studio utiliza dos grupos de seguridad para controlar el tráfico de red entre los espacios de trabajo del estudio y un clúster de Amazon EMR adjunto que se ejecuta en Amazon: EC2

- Un grupo de seguridad del motor que utiliza el puerto 18888 para comunicarse con un clúster de Amazon EMR adjunto que se ejecuta en Amazon. EC2
- Un grupo de seguridad de espacio de trabajo asociado a los espacios de trabajo de un estudio. Este grupo de seguridad incluye una regla HTTPS saliente para permitir que el espacio de trabajo enrute el tráfico a internet y debe permitir el tráfico saliente a internet en el puerto 443 para poder vincular los repositorios de Git a un espacio de trabajo.

EMR Studio usa estos grupos de seguridad además de cualquier grupo de seguridad asociado a un clúster de EMR conectado a un espacio de trabajo.

Debe crear estos grupos de seguridad cuando utilice el AWS CLI para crear un Studio.

> **Note**
>
> Puede personalizar los grupos de seguridad de EMR Studio con reglas adaptadas a su entorno, pero debe incluir las reglas que se indican en esta página. El grupo de seguridad del espacio de trabajo no puede permitir ningún tráfico entrante y el grupo de seguridad del motor debe permitir el tráfico entrante del grupo de seguridad del espacio de trabajo.

Usar los grupos de seguridad predeterminados de EMR Studio

Al utilizar la consola de Amazon EMR, puede elegir los siguientes grupos de seguridad predeterminados. EMR Studio crea los grupos de seguridad predeterminados en su nombre, los cuales contienen las reglas de entrada y salida mínimas necesarias para los espacios de trabajo de un EMR Studio.

- `DefaultEngineSecurityGroup`
- `DefaultWorkspaceSecurityGroupGit` o `DefaultWorkspaceSecurityGroupWithoutGit`

Requisitos previos

Para crear los grupos de seguridad de EMR Studio, necesita una Amazon Virtual Private Cloud (VPC) para el estudio. Debe elegir esta VPC al crear los grupos de seguridad. Debe ser la misma VPC que especifique al crear el estudio. Si planea usar Amazon EMR en EKS con EMR Studio, seleccione la VPC para los nodos de trabajo del clúster de Amazon EKS.

Instrucciones

Siga las instrucciones de Creación de un grupo de seguridad de la Guía del EC2 usuario de Amazon para instancias de Linux para crear un grupo de seguridad de motor y un grupo de seguridad de Workspace en su VPC. Los grupos de seguridad deben incluir las reglas que se resumen en las siguientes tablas.

Al crear grupos de seguridad para EMR Studio, tome nota de IDs los dos. Al crear un estudio, debe especificar cada grupo de seguridad por ID.

Grupo de seguridad del motor

EMR Studio usa el puerto 18888 para comunicarse con un clúster asociado.

Reglas de entrada

Tipo	Protocolo	Puerto	Destino	Descripción
TCP	TCP	18888	Su grupo de seguridad del espacio de trabajo de EMR Studio.	Permita el tráfico desde cualquier recurso del grupo de seguridad del espacio de trabajo para EMR Studio.

Grupo de seguridad del espacio de trabajo

Este grupo de seguridad está asociado a los espacios de trabajo de un EMR Studio.

Reglas de salida

Tipo	Protocolo	Puerto	Destino	Descripción
TCP	TCP	18888	Su grupo de seguridad del motor de EMR Studio.	Permita el tráfico a cualquier recurso del grupo de seguridad del motor para EMR Studio.
HTTPS	TCP	443	0.0.0.0/0	Permita que el tráfico de internet vincule los repositorios de Git alojados públicamente con los espacios de trabajo.

Cree AWS CloudFormation plantillas para Amazon EMR Studio

Acerca de las plantillas de clúster de EMR Studio

Puede crear AWS CloudFormation plantillas para ayudar a los usuarios de EMR Studio a lanzar nuevos clústeres de Amazon EMR en un espacio de trabajo. CloudFormation las plantillas son archivos de texto formateados en JSON o YAML. En una plantilla, describes una pila de AWS

recursos y explicas CloudFormation cómo aprovisionar esos recursos por ti. En el caso de EMR Studio, puede crear una o más plantillas que describan un clúster de Amazon EMR.

Las plantillas se organizan en AWS Service Catalog. AWS Service Catalog le permite crear y gestionar los servicios de TI más utilizados, denominados productos on AWS. Debe recopilar sus plantillas como productos en una cartera que comparta con los usuarios de EMR Studio. Tras crear las plantillas de clúster, los usuarios de Studio pueden lanzar un nuevo clúster para un espacio de trabajo con una de sus plantillas. Los usuarios deben tener permiso para crear nuevos clústeres a partir de plantillas. Puede configurar los permisos de usuario en las políticas de permisos de EMR Studio.

Para obtener más información sobre CloudFormation las plantillas, consulte Plantillas en la Guía del AWS CloudFormation usuario. Para obtener más información AWS Service Catalog, consulte Qué es AWS Service Catalog.

En el siguiente video se muestra cómo configurar plantillas de clúster en AWS Service Catalog para EMR Studio. También puede obtener más información en la entrada del blog Crear un entorno de autoservicio para cada línea de negocio mediante Amazon EMR y Service Catalog.

Parámetros de plantilla opcionales

Puede incluir opciones adicionales en la sección `Parameters` de la plantilla. Los parámetros permiten a los usuarios de Studio introducir o seleccionar valores personalizados para un clúster. Por ejemplo, puede agregar un parámetro que permita a los usuarios seleccionar una versión concreta de Amazon EMR. Para obtener más información, consulte Parámetros en la Guía del usuario de AWS CloudFormation .

En la siguiente sección `Parameters` de ejemplo se definen parámetros de entrada adicionales, como `ClusterName`, la versión de `EmrRelease` y `ClusterInstanceType`.

```
Parameters:
  ClusterName:
    Type: "String"
    Default: "Cluster_Name_Placeholder"
  EmrRelease:
    Type: "String"
    Default: "emr-6.2.0"
    AllowedValues:
    - "emr-6.2.0"
    - "emr-5.32.0"
  ClusterInstanceType:
```

```
    Type: "String"
    Default: "m5.xlarge"
    AllowedValues:
    - "m5.xlarge"
    - "m5.2xlarge"
```

Al agregar parámetros, los usuarios de Studio ven opciones de formulario adicionales tras seleccionar una plantilla de clúster. La siguiente imagen muestra opciones de formulario adicionales para EmrReleasela versión ClusterName, y InstanceType.

Requisitos previos

Antes de crear una plantilla de clúster, asegúrese de tener permisos de IAM para acceder a la vista de la consola de administrador de Service Catalog. También necesita los permisos de IAM necesarios para realizar las tareas administrativas de Service Catalog. Para obtener más información, consulte Conceder permisos a los administradores de Service Catalog.

Crear plantillas de clúster EMR

Para crear plantillas de clústeres de EMR mediante Service Catalog

1. Cree una o más CloudFormation plantillas. Puede elegir el lugar donde almacena sus plantillas. Como las plantillas son archivos de texto con formato, puede cargarlas en Amazon S3 o guardarlas en su sistema de archivos local. Para obtener más información sobre CloudFormation las plantillas, consulte Plantillas en la Guía del AWS CloudFormation usuario.

   Use las siguientes reglas para asignar un nombre a sus plantillas o compare sus nombres con el patrón `[a-zA-Z0-9][a-zA-Z0-9._-]*`.

   - El nombre de las plantillas debe comenzar por una letra o un número.
   - Los nombres de plantillas solo pueden constar de letras, números, puntos (.), guiones bajos (_) y guiones (-).

   Cada plantilla de clúster que cree debe incluir las siguientes opciones:

   Parámetros de entrada

   - ClusterName — Un nombre para el clúster que ayude a los usuarios a identificarlo una vez aprovisionado.

Salida

- `ClusterId`: el ID del clúster de EMR recién aprovisionado.

A continuación, se muestra un ejemplo AWS CloudFormation de plantilla en formato YAML para un clúster con dos nodos. En la plantilla de ejemplo se incluyen las opciones de plantilla necesarias y se definen parámetros de entrada adicionales para `EmrRelease` y `ClusterInstanceType`.

```
awsTemplateFormatVersion: 2010-09-09

Parameters:
  ClusterName:
    Type: "String"
    Default: "Example_Two_Node_Cluster"
  EmrRelease:
    Type: "String"
    Default: "emr-6.2.0"
    AllowedValues:
    - "emr-6.2.0"
    - "emr-5.32.0"
  ClusterInstanceType:
    Type: "String"
    Default: "m5.xlarge"
    AllowedValues:
    - "m5.xlarge"
    - "m5.2xlarge"

Resources:
  EmrCluster:
    Type: AWS::EMR::Cluster
    Properties:
      Applications:
      - Name: Spark
      - Name: Livy
      - Name: JupyterEnterpriseGateway
      - Name: Hive
      EbsRootVolumeSize: '10'
      Name: !Ref ClusterName
      JobFlowRole: EMR_EC2_DefaultRole
```

```
      ServiceRole: EMR_DefaultRole_V2
      ReleaseLabel: !Ref EmrRelease
      VisibleToAllUsers: true
      LogUri:
        Fn::Sub: 's3://aws-logs-${AWS::AccountId}-${AWS::Region}/elasticmapreduce/'
      Instances:
        TerminationProtected: false
        Ec2SubnetId: 'subnet-ab12345c'
        MasterInstanceGroup:
          InstanceCount: 1
          InstanceType: !Ref ClusterInstanceType
        CoreInstanceGroup:
          InstanceCount: 1
          InstanceType: !Ref ClusterInstanceType
          Market: ON_DEMAND
          Name: Core

Outputs:
  ClusterId:
    Value:
      Ref: EmrCluster
    Description: The ID of the  EMR cluster
```

2. Crea una cartera para tus plantillas de clúster en la misma AWS cuenta que tu Studio.
   a. Abre la AWS Service Catalog consola en https://console.aws.amazon.com/servicecatalog/.
   b. En el menú de navegación izquierdo, seleccione Carteras.
   c. En la página Crear cartera, introduzca la información solicitada.
   d. Elija Crear. AWS Service Catalog crea la cartera y muestra los detalles de la cartera.
3. Utilice los siguientes pasos para agregar plantillas de clústeres como productos de AWS Service Catalog .
   a. Navegue a la página Productos en Administración en la consola de administración de AWS Service Catalog .
   b. Seleccione Cargar un producto nuevo.
   c. Introduzca el nombre del producto y el propietario.
   d. Especifique el archivo de plantilla en Detalles de la versión.
   e. Seleccione Revisar para revisar la configuración del producto y, a continuación, seleccione Crear producto.

4. Complete los siguientes pasos para agregar sus productos a su cartera.
   a. Navegue a la página Productos en la consola de administración de AWS Service Catalog .
   b. Seleccione su producto, seleccione Acciones y, a continuación, seleccione Agregar producto a la cartera.
   c. Seleccione su cartera y, a continuación, seleccione Agregar producto a la cartera.
5. Cree una restricción de lanzamiento para sus productos. Una restricción de lanzamiento es un rol de IAM que especifica los permisos de los usuarios para lanzar un producto. Puede personalizar sus restricciones de lanzamiento, pero debe permitir permisos de uso CloudFormation, Amazon EMR y. AWS Service Catalog Para obtener más información e instrucciones, consulte Restricciones de lanzamiento de Service Catalog.
6. Aplique su restricción de lanzamiento a cada producto de su cartera. Debe aplicar la restricción de lanzamiento a cada producto de forma individual.
   a. Seleccione su cartera en la página Carteras de la consola de administración de AWS Service Catalog .
   b. Elija la pestaña Constraints (Restricciones) y elija Create constraint (Crear restricción).
   c. Seleccione su producto y, en Tipo de restricción, Lanzamiento. Elija Continuar.
   d. Seleccione su rol de restricción de lanzamiento en la sección Restricción de lanzamiento y, a continuación, seleccione Crear.
7. Concede acceso a su cartera.
   a. Seleccione su cartera en la página Carteras de la consola de administración de AWS Service Catalog .
   b. Amplíe la pestaña Grupos, roles y usuarios y seleccione Agregar grupos, roles y usuarios.
   c. Busque su rol de IAM de EMR Studio en la pestaña Roles, seleccione su rol y elija Agregar acceso.

Si usa…	Conceder acceso a…
Autenticación de IAM	Sus usuarios nativos
Federación de IAM	Su rol de IAM en la federación
Federación de IAM Identity Center	Su rol de usuario de EMR Studio

Establecer el acceso y los permisos para los repositorios basados en Git

EMR Studio admite los siguientes servicios basados en Git:

- AWS CodeCommit
- GitHub
- Bitbucket
- GitLab

Para permitir que los usuarios de EMR Studio asocien un repositorio de Git a un espacio de trabajo, configure los siguientes requisitos de acceso y permisos. También puede configurar los repositorios basados en Git que aloje en una red privada siguiendo las instrucciones que se indican en Configurar un repositorio de Git alojado de forma privada para EMR Studio.

Acceso a internet en clúster

Tanto los clústeres de Amazon EMR que se ejecutan en EC2 Amazon como los de Amazon EMR en los clústeres EKS conectados a Studio Workspaces deben estar en una subred privada que utilice una puerta de enlace de traducción de direcciones de red (NAT) o deben poder acceder a Internet a través de una puerta de enlace privada virtual. Para obtener más información, consulte Opciones de Amazon VPC al lanzar un clúster.

Los grupos de seguridad que utilice con EMR Studio también deben incluir una regla de salida que permita que los espacios de trabajo enruten el tráfico a internet desde un clúster de EMR asociado. Para obtener más información, consulte Definir grupos de seguridad para controlar el tráfico de red de EMR Studio.

> **Important**
>
> Si la interfaz de red se encuentra en una subred pública, no podrá comunicarse con internet a través de una puerta de enlace de Internet (IGW).

Permisos para AWS Secrets Manager

Para permitir que los usuarios de EMR Studio accedan a los repositorios de Git con secretos almacenados en AWS Secrets Manager, agregue una política de permisos al rol de servicio de EMR Studio que permite la operación `secretsmanager:GetSecretValue`.

Para obtener información sobre cómo vincular los repositorios basados en Git a los espacios de trabajo, consulte Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio.

Configurar un repositorio de Git alojado de forma privada para EMR Studio

Siga las instrucciones siguientes para configurar los repositorios alojados de forma privada para Amazon EMR Studio. Proporcione un archivo de configuración con información sobre sus servidores DNS y Git. EMR Studio utiliza esta información para configurar los espacios de trabajo que pueden enrutar el tráfico a sus repositorios autoadministrados.

> **Note**
>
> Si configura `DnsServerIpV4`, EMR Studio utilizará su servidor DNS para resolver tanto su `GitServerDnsName` como su punto de conexión de Amazon EMR, como `elasticmapreduce.us-east-1.amazonaws.com`. Para configurar un punto de conexión para Amazon EMR, conéctese a su punto de conexión a través de la VPC que utilice con su estudio. De este modo, se garantiza que el punto de conexión de Amazon EMR se especifique en una IP privada. Para obtener más información, consulte Conexión a Amazon EMR mediante un punto de conexión de VPC de tipo interfaz.

Requisitos previos

Antes de configurar un repositorio de Git alojado de forma privada para EMR Studio, necesita una ubicación de almacenamiento de Amazon S3 en la que EMR Studio pueda realizar copias de seguridad de los espacios de trabajo y los archivos de cuadernos del estudio. Utilice el mismo bucket de S3 que especifique al crear un estudio.

Para configurar un repositorio de Git alojado de forma privada para EMR Studio

1. Cree un archivo de configuración mediante la siguiente plantilla. Incluya los siguientes valores para cada servidor de Git que desee especificar en la configuración:

   - **`DnsServerIpV4`**- La IPv4 dirección de su servidor DNS. Si proporciona valores para `DnsServerIpV4` y `GitServerIpV4List`, el valor de `DnsServerIpV4` tiene prioridad y EMR Studio utiliza `DnsServerIpV4` para resolver el `GitServerDnsName`.

> **Note**
>
> Para usar repositorios de Git alojados de forma privada, su servidor DNS debe permitir el acceso entrante desde EMR Studio. Le instamos a que proteja su servidor DNS contra otros accesos no autorizados.

- **`GitServerDnsName`**: el nombre de DNS del servidor de Git. Por ejemplo, `"git.example.com"`.
- **`GitServerIpV4List`**- Una lista de IPv4 direcciones que pertenecen a tus servidores Git.

```
[
    {
        "Type": "PrivatelyHostedGitConfig",
        "Value": [
            {
                "DnsServerIpV4": "<10.24.34.xxx>",
                "GitServerDnsName": "<enterprise.git.com>",
                "GitServerIpV4List": [
                    "<xxx.xxx.xxx.xxx>",
                    "<xxx.xxx.xxx.xxx>"
                ]
            },
            {
                "DnsServerIpV4": "<10.24.34.xxx>",
                "GitServerDnsName": "<git.example.com>",
                "GitServerIpV4List": [
                    "<xxx.xxx.xxx.xxx>",
                    "<xxx.xxx.xxx.xxx>"
                ]
            }
        ]
    }
]
```

2. Guarde su archivo de configuración como `configuration.json`.
3. Cargue el archivo de configuración en su ubicación de almacenamiento predeterminada de Amazon S3, en una carpeta llamada`life-cycle-configuration`. Por ejemplo, si su ubicación de S3 predeterminada es `s3://`*`amzn-s3-demo-bucket`*`/workspace`, el archivo

de configuración estaría en `s3://amzn-s3-demo-bucket/workspace/life-cycle-configuration/configuration.json`.

> **Important**
>
> Le instamos a que restrinja el acceso a su carpeta `life-cycle-configuration` a los administradores de Studio y a su rol de servicio de EMR Studio, y a que proteja `configuration.json` contra el acceso no autorizado. Para obtener instrucciones, consulte Controlar el acceso a un bucket con políticas de usuario o Prácticas recomendadas de seguridad para Amazon S3.

Para ver las instrucciones de carga, consulte Creación de una carpeta y Carga de objetos en la Guía del usuario de Amazon Simple Storage Service. Para aplicar la configuración a un espacio de trabajo existente, ciérrelo y reinícielo después de cargar el archivo de configuración en Amazon S3.

Optimizar los trabajos de Spark en EMR Studio

Al ejecutar un trabajo de Spark con EMR Studio, hay algunos pasos que puede seguir para asegurarse de que optimiza los recursos de su clúster de Amazon EMR.

Prolongar la sesión de Livy

Si utiliza Apache Livy junto con Spark en su clúster de Amazon EMR, le recomendamos que aumente el tiempo de espera de la sesión de Livy de la siguiente manera:

- Al crear un clúster de Amazon EMR, defina esta clasificación de configuración en el campo Introducir configuración.

```
[
    {
        "Classification": "livy-conf",
        "Properties": {
          "livy.server.session.timeout": "8h"
        }
    }
]
```

- En el caso de un clúster de EMR que ya se esté ejecutando, conéctese a su clúster mediante `ssh` y establezca la clasificación de configuración `livy-conf` en `/etc/livy/conf/livy.conf`.

```
[
    {
        "Classification": "livy-conf",
        "Properties": {
          "livy.server.session.timeout": "8h"
        }
    }
]
```

  Es posible que necesite reiniciar Livy después de cambiar la configuración.

- Si no quiere que se agote el tiempo de espera de su sesión de Livy, configure la propiedad `livy.server.session.timeout-check` como `false` en `/etc/livy/conf/livy.conf`.

Ejecutar Spark en el modo de clúster

En el modo de clúster, el controlador de Spark se ejecuta en un nodo de núcleo en lugar de en el nodo principal, lo que mejora la utilización de los recursos en el nodo principal.

Para ejecutar su aplicación de Spark en el modo de clúster en lugar del modo cliente predeterminado, elija el modo Clúster al configurar el Modo de implementación mientras configura el paso de Spark en el nuevo clúster de Amazon EMR. Para obtener más información, consulte Cluster mode overview en la documentación de Apache Spark.

Aumentar la memoria del controlador de Spark

Para aumentar la memoria del controlador de Spark, configure su sesión de Spark con el comando mágico `%%configure` de su cuaderno de EMR, como en el siguiente ejemplo.

```
%%configure -f
{"driverMemory": "6000M"}
```

Uso de un Amazon EMR Studio

Esta sección contiene temas que lo ayudan a configurar un Amazon EMR Studio e interactuar con él.

En el siguiente video, se incluye información práctica, como la creación de un nuevo espacio de trabajo y el lanzamiento de un nuevo clúster de Amazon EMR con una plantilla de clúster. El video también muestra un cuaderno de muestra.

En esta sección se incluyen los temas siguientes, que le ayudarán a trabajar en un EMR Studio:

- Conozca los espacios de trabajo de EMR Studio
- Configuración de la colaboración en el espacio de trabajo en EMR Studio
- Ejecutar un espacio de trabajo de EMR Studio con un rol de tiempo de ejecución
- Ejecutar los cuadernos de espacios de trabajo de Amazon EMR Studio mediante programación
- Examinar los datos con SQL Explorer para EMR Studio
- Asociar computación a un espacio de trabajo de EMR Studio
- Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio
- Uso del editor SQL de Amazon Athena en EMR Studio
- CodeWhisperer Integración de Amazon con EMR Studio Workspaces
- Depurar aplicaciones y trabajos con EMR Studio
- Instalar kernels y bibliotecas en un espacio de trabajo de EMR Studio
- Mejorar los kernels con comandos magic en EMR Studio
- Usar cuadernos multilingües con kernels de Spark

Conozca los espacios de trabajo de EMR Studio

Cuando utiliza un EMR Studio, puede crear y configurar diferentes espacios de trabajo para organizar y ejecutar cuadernos. En esta sección se describe la creación de espacios de trabajo y el trabajo con ellos. Para obtener información general conceptual, consulte Workspaces en la página Cómo funciona Amazon EMR Studio.

En esta sección se cubren los temas siguientes, que le ayudarán a usar espacios de trabajo de EMR Studio:

- Crear un espacio de trabajo de EMR Studio
- Inicie un espacio de trabajo en EMR Studio
- Conocer la interfaz de usuario del espacio de trabajo en EMR Studio
- Examinar los ejemplos de libretas en un espacio de trabajo de EMR Studio
- Guardar contenido del espacio de trabajo en EMR Studio

- Eliminar un espacio de trabajo y archivos de cuaderno en EMR Studio
- Explicación del estado del espacio de trabajo
- Resolver problemas de conectividad con el espacio de trabajo

Crear un espacio de trabajo de EMR Studio

Puede crear espacios de trabajo de EMR Studio para ejecutar código de cuadernos mediante la interfaz de EMR Studio.

Para crear un espacio de trabajo en un EMR Studio

1. Inicie sesión en su EMR Studio.
2. Seleccione Crear un espacio de trabajo.
3. Ingrese un Nombre de espacio de trabajo y una Description. Asignar un nombre a un espacio de trabajo le ayuda a identificarlo en la página Espacios de trabajo.
4. Si quiere trabajar con otros usuarios de Studio en este espacio de trabajo en tiempo real, active la colaboración en el espacio de trabajo. Puede configurar los colaboradores después de lanzar el espacio de trabajo.
5. Si quiere asociar un clúster a un espacio de trabajo, expanda la sección Configuración avanzada. Si lo prefiere, puede asociar un clúster más adelante. Para obtener más información, consulte Asociar computación a un espacio de trabajo de EMR Studio.

   > **Note**
   >
   > Para aprovisionar un clúster nuevo, necesita los permisos de acceso de su administrador.

   Elija una de las opciones de clúster para el espacio de trabajo y asocie el clúster. Para obtener más información sobre el aprovisionamiento de un clúster al crear un espacio de trabajo, consulte Crear y asociar un nuevo clúster de EMR a un espacio de trabajo de EMR Studio.
6. Seleccione Crear un espacio de trabajo en la parte inferior derecha de la página.

Tras crear un espacio de trabajo, EMR Studio abrirá la página Espacios de trabajo. En la parte superior de la página, verá un banner verde indicando que la operación se ha realizado correctamente y encontrará el espacio de trabajo recién creado en la lista.

De forma predeterminada, los espacios de trabajo son compartidos y pueden verlos todos los usuarios de Studio. Sin embargo, solo un usuario puede abrir y trabajar en un espacio de trabajo a la vez. Para trabajar simultáneamente con otros usuarios, puede Configuración de la colaboración en el espacio de trabajo en EMR Studio.

Inicie un espacio de trabajo en EMR Studio

Para empezar a trabajar con los archivos de cuaderno, abra un espacio de trabajo para acceder al editor del cuaderno. La página Espacios de trabajo de un estudio muestra todos los espacios de trabajo a los que tiene acceso, con detalles como el nombre, el estado, la hora de creación y la última modificación.

> **Note**
>
> Si tenía cuadernos de EMR en la antigua consola de Amazon EMR, los encontrará en la consola como espacios de trabajo de EMR Studio. Los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales para acceder o crear espacios de trabajo. Si ha creado recientemente un cuaderno en la consola anterior, puede que tenga que actualizar la lista de espacios de trabajo para poder verlo en la consola. Para obtener más información sobre la transición, consulte Cuadernos de Amazon EMR está disponible como Espacios de trabajo de Amazon EMR Studio en la consola y Administración de clústeres de Amazon EMR con la consola.

Para iniciar un espacio de trabajo para editar y ejecutar cuadernos

1. En la página Espacios de trabajo de su estudio, busque el espacio de trabajo. Puede filtrar la lista por palabra clave o por valor de columna.
2. Elija el nombre del espacio de trabajo para iniciarlo en una nueva pestaña del navegador. Si está inactivo, es posible que el espacio de trabajo tarde unos minutos en abrirse. Como alternativa, seleccione la fila del espacio de trabajo y, a continuación, seleccione Iniciar espacio de trabajo. Puede elegir entre las siguientes opciones:
   - Inicio rápido: inicie rápidamente su espacio de trabajo con las opciones predeterminadas. Seleccione Inicio rápido si desea adjuntar clústeres al espacio de trabajo JupyterLab.
   - Inicio con opciones: inicie su espacio de trabajo con opciones personalizadas. Puede elegir lanzarlo en Jupyter o JupyterLab adjuntar su espacio de trabajo a un clúster de EMR y seleccionar sus grupos de seguridad.

> **Note**
>
> Solo un usuario puede abrir y trabajar en un espacio de trabajo a la vez. Si selecciona un espacio de trabajo que ya está en uso, EMR Studio mostrará una notificación cuando intente abrirlo. La columna Usuario de la página Espacios de trabajo muestra el usuario que trabaja en el espacio de trabajo.

Conocer la interfaz de usuario del espacio de trabajo en EMR Studio

La interfaz de usuario de EMR Studio Workspace se basa en la JupyterLabinterfaz con pestañas marcadas por iconos en la barra lateral izquierda. Al hacer una pausa sobre un icono, aparece información sobre herramientas que muestra el nombre de la pestaña. Seleccione las pestañas de la barra lateral izquierda para acceder a los siguientes paneles.

- Explorador de archivos: muestra los archivos y directorios del espacio de trabajo, así como los archivos y directorios de los repositorios de Git enlazados.
- Ejecución de kernels y terminales: muestra todos los kernels y terminales que se ejecutan en el espacio de trabajo. Para obtener más información, consulte Administración de núcleos y terminales en la documentación oficial. JupyterLab
- Git: proporciona una interfaz gráfica de usuario para ejecutar comandos en los repositorios de Git asociados al espacio de trabajo. Este panel es una JupyterLab extensión llamada jupyterlab-git. Para obtener más información, consulte jupyterlab-git.
- Clústeres de EMR: le permite asociar un clúster al espacio de trabajo o separarlo a fin de ejecutar el código del cuaderno. El panel de configuración del clúster de EMR también ofrece opciones de configuración avanzadas para ayudarle a crear y asociar un clúster nuevo al espacio de trabajo. Para obtener más información, consulte Crear y asociar un nuevo clúster de EMR a un espacio de trabajo de EMR Studio.
- Repositorio de Git de Amazon EMR: le ayuda a vincular el espacio de trabajo con hasta tres repositorios de Git. Para obtener información más detallada e instrucciones, consulte Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio.
- Ejemplos de cuadernos: proporciona una lista de ejemplos de cuadernos que puede guardar en el espacio de trabajo. También puede acceder a los ejemplos seleccionando Ejemplos de cuadernos en la página de inicio del espacio de trabajo.

- Comandos: ofrece una forma de buscar y ejecutar comandos mediante el teclado. JupyterLab Para obtener más información, consulte la página de la paleta de comandos de la documentación. JupyterLab
- Herramientas de cuaderno: le permiten seleccionar y configurar opciones como el tipo de diapositiva de la celda y los metadatos. La opción Herramientas del cuaderno aparece en la barra lateral izquierda después de abrir un archivo de cuaderno.
- Pestañas abiertas: muestra los documentos y las actividades abiertos en el área de trabajo principal para que pueda ir directamente a una pestaña abierta. Para obtener más información, consulte la página de pestañas y modo de documento único en la JupyterLab documentación.
- Colaboración: le permite activar o desactivar la colaboración en el espacio de trabajo y administrar a los colaboradores. Para ver el panel Colaboración, debe tener los permisos necesarios. Para obtener más información, consulte Establecer la propiedad de la colaboración en el espacio de trabajo.

Examinar los ejemplos de libretas en un espacio de trabajo de EMR Studio

Cada espacio de trabajo de EMR Studio incluye un conjunto de ejemplos de cuadernos que puede utilizar para explorar las características de EMR Studio. Para editar o ejecutar un ejemplo de cuaderno, puede guardarlo en el espacio de trabajo.

Para guardar un ejemplo de cuaderno en un espacio de trabajo

1. En la barra lateral izquierda, seleccione la pestaña Ejemplos de cuadernos para abrir el panel Ejemplos de cuadernos. También puede acceder a los ejemplos seleccionando Ejemplos de cuadernos en la página de inicio del espacio de trabajo.
2. Elija un ejemplo de cuaderno para ver su vista previa en el área de trabajo principal. El ejemplo es de solo lectura.
3. Para guardar el ejemplo del cuaderno en el espacio de trabajo, seleccione Guardar en el espacio de trabajo. EMR Studio guarda el ejemplo en su directorio de inicio. Tras guardar un ejemplo de cuaderno en el espacio de trabajo, puede cambiar su nombre, editarlo y ejecutarlo.

Para obtener más información sobre los ejemplos de cuadernos, consulte el repositorio de ejemplos GitHub de cuadernos de EMR Studio.

Guardar contenido del espacio de trabajo en EMR Studio

Cuando trabaja en el editor del cuaderno de un espacio de trabajo, EMR Studio guarda el contenido de las celdas del cuaderno y la salida para usted en la ubicación de Amazon S3 asociada al estudio. Este proceso de copia de seguridad preserva el trabajo entre sesiones.

También puede guardar un cuaderno pulsando CTRL+S en la pestaña del cuaderno abierto o utilizando una de las opciones de guardado de la sección Archivo.

Otra forma de hacer copias de seguridad de los archivos de cuaderno en un espacio de trabajo consiste en asociar el espacio de trabajo a un repositorio basado en Git y sincronizar los cambios con el repositorio remoto. De este modo, también podrá guardar y compartir cuadernos con los miembros del equipo que usen un espacio de trabajo o un estudio diferente. Para obtener instrucciones, consulte Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio.

Eliminar un espacio de trabajo y archivos de cuaderno en EMR Studio

Al eliminar un archivo de cuaderno de un espacio de trabajo de EMR Studio, se elimina el archivo del Explorador de archivos y EMR Studio elimina su copia de seguridad en Amazon S3. No tiene que tomar ninguna otra medida para evitar los cargos de almacenamiento cuando elimina un archivo de un espacio de trabajo.

Al eliminar un espacio de trabajo completo, los archivos de cuaderno y las carpetas permanecerán en la ubicación de almacenamiento de Amazon S3. Los archivos siguen acumulando gastos de almacenamiento. Para evitar gastos de almacenamiento, elimine de Amazon S3 todos los archivos y carpetas de los que se haya hecho una copia de seguridad y que estén asociados a su espacio de trabajo eliminado.

Para eliminar un archivo de cuaderno de un espacio de trabajo de EMR Studio

1. Seleccione el panel del Explorador de archivos en la barra lateral izquierda del espacio de trabajo.
2. Seleccione el archivo o la carpeta que desea eliminar. Haga clic con el botón derecho en la selección y seleccione Eliminar. El archivo desaparecerá de la lista. EMR Studio elimina automáticamente el archivo o la carpeta de Amazon S3.

From the Workspace UI

Eliminar un espacio de trabajo y sus archivos de copia de seguridad asociados de EMR Studio

1. Inicie sesión en su EMR Studio con la URL de acceso de su estudio y elija Espacios de trabajo en el menú de navegación de la izquierda.
2. Busque su espacio de trabajo en la lista y seleccione la casilla de verificación que aparece junto a su nombre. Puede seleccionar varios espacios de trabajo para eliminarlos al mismo tiempo.
3. Seleccione Eliminar en la esquina superior derecha de la lista de Espacios de trabajo y confirme que desea eliminar los espacios de trabajo seleccionados. Elija Eliminar para confirmar.
4. Si desea eliminar los archivos de cuadernos que estaban asociados al espacio de trabajo eliminado de Amazon S3, siga las instrucciones para eliminar objetos en la Guía del usuario de la consola de Amazon Simple Storage Service. Si usted no creó el estudio, consulte a su administrador del estudio para determinar la ubicación de la copia de seguridad de Amazon S3 para el espacio de trabajo eliminado.

From the Workspaces list

Eliminar un espacio de trabajo y sus archivos de copia de seguridad asociados de la lista de espacios de trabajo

1. Navegue hasta la lista Espacios de trabajo en la consola.
2. Seleccione el espacio de trabajo que desee eliminar de la lista y, a continuación, seleccione Acciones.
3. Elija Eliminar.
4. Si desea eliminar los archivos de cuadernos que estaban asociados al espacio de trabajo eliminado de Amazon S3, siga las instrucciones para eliminar objetos en la Guía del usuario de la consola de Amazon Simple Storage Service. Si usted no creó el estudio, consulte a su administrador del estudio para determinar la ubicación de la copia de seguridad de Amazon S3 para el espacio de trabajo eliminado.

Explicación del estado del espacio de trabajo

Tras crear un espacio de trabajo de EMR Studio, aparece como una fila en la lista Espacios de trabajo de su estudio con su nombre, estado, hora de creación y fecha de la última modificación. En la tabla siguiente se describen los estados de los espacios de trabajo.

Estado	Descripción
Iniciando	El espacio de trabajo se está preparando, pero aún no está listo para usarse. No puede abrir un espacio de trabajo cuando se encuentra en el estado Iniciando.
Ready	Puede abrir el espacio de trabajo para usar el editor de cuadernos, pero debe asociar el espacio de trabajo a un clúster de EMR antes de poder ejecutar el código del cuaderno.
Asociando	El espacio de trabajo se está asociando a un clúster.
Attached (Asociado)	El espacio de trabajo se ha asociado a un clúster de EMR y está listo para que escriba y ejecute el código del cuaderno. Si el estado de un espacio de trabajo no es Asociado, debe asociarlo a un clúster para poder ejecutar el código del cuaderno.
Inactivo	El espacio de trabajo se ha detenido. Para reactivar un espacio de trabajo inactivo, selecciónelo en la lista de espacios de trabajo. El estado cambia de Inactivo a Empezando y a Preparado al seleccionar el espacio de trabajo.
Deteniéndose	El espacio de trabajo se está apagando y pasará a estar Inactivo. Al detener un espacio de trabajo, se cancelan los kernels de cuadernos correspondientes. EMR Studio

Estado	Descripción
	detiene los cuadernos que han estado inactivos durante mucho tiempo.
Eliminando	Al eliminar un espacio de trabajo, EMR Studio lo marca para su eliminación e inicia el proceso de eliminación. Una vez finalizado el proceso de eliminación, el espacio de trabajo desaparec e de la lista. Al eliminar un espacio de trabajo, los archivos de cuaderno permanecerán en la ubicación de almacenamiento de Amazon S3.

Resolver problemas de conectividad con el espacio de trabajo

Para resolver los problemas de conectividad de un espacio de trabajo, puede detenerlo y reiniciarlo. Al reiniciar un espacio de trabajo, EMR Studio lanza el espacio de trabajo en una zona de disponibilidad diferente o en una subred diferente que esté asociada a su estudio.

Para detener y reiniciar un espacio de trabajo de EMR Studio

1. Cierre el espacio de trabajo en su navegador.
2. Navegue hasta la lista Espacios de trabajo en la consola.
3. En la lista, seleccione su espacio de trabajo y, a continuación, seleccione Acciones.
4. Seleccione Detener y espera a que el estado del espacio de trabajo cambie de Deteniendo a Inactivo.
5. Vuelva a seleccionar Acciones y, a continuación, seleccione Iniciar para reiniciar el espacio de trabajo.
6. Espere a que el estado del espacio de trabajo cambie de Empezando a Listo y, a continuación, elija el nombre del espacio de trabajo para volver a abrirlo en una nueva pestaña del navegador.

Configuración de la colaboración en el espacio de trabajo en EMR Studio

La colaboración en el espacio de trabajo le permite escribir y ejecutar código de cuadernos de forma simultánea con otros miembros del equipo. Cuando trabaje en el mismo archivo de cuaderno, verá

los cambios a medida que los hagan los colaboradores. Puede habilitar la colaboración al crear un espacio de trabajo o activar y desactivar la colaboración en un espacio de trabajo existente.

> **Note**
>
> La colaboración en el espacio de trabajo de EMR Studio no es compatible con las aplicaciones interactivas de EMR sin servidor o si la propagación de identidades de confianza está habilitada.

Requisitos previos

Antes de configurar la colaboración en un espacio de trabajo, asegúrese de completar las siguientes tareas:

- Asegúrese de que el administrador de EMR Studio le haya otorgado los permisos necesarios. Por ejemplo, la siguiente instrucción permite a un usuario configurar la colaboración para cualquier espacio de trabajo con la clave de etiqueta `creatorUserId` cuyo valor coincida con el ID del usuario (indicado por la variable de política `aws:userId`).

```
{
    "Sid": "UserRolePermissionsForCollaboration",
    "Action": [
        "elasticmapreduce:UpdateEditor",
        "elasticmapreduce:PutWorkspaceAccess",
        "elasticmapreduce:DeleteWorkspaceAccess",
        "elasticmapreduce:ListWorkspaceAccessIdentities"
    ],
    "Resource": "*",
    "Effect": "Allow",
    "Condition": {
        "StringEquals": {
            "elasticmapreduce:ResourceTag/creatorUserId": "${aws:userid}"
        }
    }
}
```

- Asegúrese de que el rol de servicio asociado a su EMR Studio tenga los permisos necesarios para habilitar y configurar la colaboración en el espacio de trabajo, como se muestra en la siguiente declaración de ejemplo.

```
{
    "Sid": "AllowWorkspaceCollaboration",
    "Effect": "Allow",
    "Action": [
        "iam:GetUser",
        "iam:GetRole",
        "iam:ListUsers",
        "iam:ListRoles",
        "sso:GetManagedApplicationInstance",
        "sso-directory:SearchUsers"
    ],
    "Resource": "*"
}
```

Para obtener más información, consulte Crear un rol de servicio de EMR Studio.

Para habilitar la colaboración en el espacio de trabajo y agregar colaboradores

1. En su espacio de trabajo, seleccione el icono de Colaboración en la pantalla de inicio o en la parte inferior del panel izquierdo.

   > **Note**
   >
   > No verá el panel Colaboración a menos que el administrador de Studio le haya dado permiso para configurar la colaboración en el espacio de trabajo. Para obtener más información, consulte Establecer la propiedad de la colaboración en el espacio de trabajo.

2. Asegúrese de que la opción Permitir la colaboración en el espacio de trabajo esté activada. Cuando habilita la colaboración, solo usted y los colaboradores que agregue podrán ver el espacio de trabajo en la lista de la página Espacios de trabajo de Studio.
3. Introduzca el nombre de un colaborador. En el espacio de trabajo se puede tener un máximo de cinco colaboradores, incluido usted. Un colaborador puede ser cualquier usuario con acceso a su EMR Studio. Si no introduce un colaborador, el espacio de trabajo es un espacio de trabajo privado al que solo usted puede acceder.

   En la siguiente tabla se especifican los valores de colaborador aplicables que se deben introducir en función del tipo de identidad del propietario.

> **Note**
>
> Un propietario solo puede invitar a colaboradores con el mismo tipo de identidad. Por ejemplo, un usuario solo puede agregar a otros usuarios, y un usuario de IAM Identity Center solo puede agregar a otros usuarios de IAM Identity Center.

Modo de autenticación	Valor que se debe introducir en Nombre del colaborador
Autenticación de IAM	un nombre de usuario. Este es el nombre que ve un usuario al iniciar sesión en la AWS Management Console.
Federación de IAM	El nombre de un rol de IAM y un nombre de sesión opcional. Para agregar todos los usuarios federados que asumen el mismo rol de IAM, especifiq ue el nombre de un rol de IAM para la federación. Para agregar un único usuario como colaborador, especifique un rol y un nombre de sesión. Por ejemplo, `MyRoleNam e:MySessionName` .
SSO	Un nombre de usuario de IAM Identity Center, como `user@example.com.` .

4. Elija Agregar. El colaborador ahora puede ver el espacio de trabajo en su página Espacios de trabajo de EMR Studio e iniciar el espacio de trabajo para usarlo en tiempo real con usted.

> **Note**
>
> Si inhabilita la colaboración en el espacio de trabajo, el espacio de trabajo vuelve a su estado compartido y todos los usuarios de Studio lo pueden ver. En el estado compartido, solo un usuario de Studio puede abrir un espacio de trabajo y trabajar en él a la vez.

Ejecutar un espacio de trabajo de EMR Studio con un rol de tiempo de ejecución

> **Note**
>
> La funcionalidad del rol en tiempo de ejecución descrita en esta página solo se aplica a Amazon EMR que se ejecuta en Amazon EC2 y no hace referencia a la funcionalidad del rol en tiempo de ejecución en las aplicaciones interactivas EMR Serverless. Para obtener más información sobre cómo utilizar los roles de tiempo de ejecución en EMR sin servidor, consulte Rol de tiempo de ejecución de tareas en la Guía del usuario de Amazon EMR sin servidor.

Un rol en tiempo de ejecución es un rol AWS Identity and Access Management (IAM) que puede especificar al enviar un trabajo o una consulta a un clúster de Amazon EMR. El trabajo o la consulta que envíe a su clúster de EMR utiliza el rol de tiempo de ejecución para acceder a AWS los recursos, como los objetos de Amazon S3.

Al adjuntar un espacio de trabajo de EMR Studio a un clúster de EMR que usa Amazon EMR 6.11 o una versión posterior, puede seleccionar un rol de tiempo de ejecución para el trabajo o la consulta que envíe para usarlo cuando acceda a los recursos. AWS Sin embargo, si el clúster de EMR no admite funciones de tiempo de ejecución, el clúster de EMR no asumirá la función cuando acceda a los recursos. AWS

Antes de poder utilizar un rol de tiempo de ejecución con un espacio de trabajo de Amazon EMR Studio, un administrador debe configurar los permisos de usuario para que el usuario de Studio pueda llamar a la API `elasticmapreduce:GetClusterSessionCredentials` en el rol de tiempo de ejecución. A continuación, lance un nuevo clúster con un rol de tiempo de ejecución que pueda usar con su espacio de trabajo de Amazon EMR Studio.

En esta página

- Configurar los permisos de usuario para el rol de tiempo de ejecución
- Lanzar un clúster nuevo con un rol de tiempo de ejecución
- Utilizar el clúster de EMR con un rol de tiempo de ejecución en los espacios de trabajo
- Consideraciones

Configurar los permisos de usuario para el rol de tiempo de ejecución

Configure los permisos de usuario para que el usuario de Studio pueda llamar a la API `elasticmapreduce:GetClusterSessionCredentials` en el rol de tiempo de ejecución que quiera usar. También debe configurar the section called "Permisos de usuario de Studio (EC2, EKS)" para que el usuario pueda empezar a usar Studio.

> **Warning**
>
> Para conceder este permiso, cree una condición basada en la clave de `elasticmapreduce:ExecutionRoleArn` contexto cuando conceda a la persona que llama acceso para llamar al. `GetClusterSessionCredentials` APIs En los siguientes ejemplos, se muestra cómo hacerlo.

```
{
      "Sid": "AllowSpecificExecRoleArn",
      "Effect": "Allow",
      "Action": [
          "elasticmapreduce:GetClusterSessionCredentials"
      ],
      "Resource": "*",
      "Condition": {
          "StringEquals": {
              "elasticmapreduce:ExecutionRoleArn": [
                  "arn:aws:iam::111122223333:role/test-emr-demo1",
                  "arn:aws:iam::111122223333:role/test-emr-demo2"
              ]
          }
      }
  }
```

En el siguiente ejemplo, se muestra cómo permitir que una entidad principal de IAM utilice un rol de IAM denominado `test-emr-demo3` como rol de tiempo de ejecución. Además, el titular de la política solo podrá acceder a los clústeres de Amazon EMR con el ID de clúster `j-123456789`.

```
{
    "Sid":"AllowSpecificExecRoleArn",
    "Effect":"Allow",
    "Action":[
        "elasticmapreduce:GetClusterSessionCredentials"
    ],
    "Resource": [
          "arn:aws:elasticmapreduce:<region>:111122223333:cluster/j-123456789"
     ],
    "Condition":{
        "StringEquals":{
            "elasticmapreduce:ExecutionRoleArn":[
                "arn:aws:iam::111122223333:role/test-emr-demo3"
            ]
        }
    }
}
```

El siguiente ejemplo permite que una entidad principal de IAM utilice cualquier rol de IAM cuyo nombre comience por la cadena `test-emr-demo4` como rol de tiempo de ejecución. Además, el titular de la política solo podrá acceder a los clústeres de Amazon EMR etiquetados con el par clave-valor `tagKey: tagValue`.

```
{
    "Sid":"AllowSpecificExecRoleArn",
    "Effect":"Allow",
    "Action":[
        "elasticmapreduce:GetClusterSessionCredentials"
    ],
    "Resource": "*",
    "Condition":{
        "StringEquals":{
             "elasticmapreduce:ResourceTag/tagKey": "tagValue"
        },
        "StringLike":{
            "elasticmapreduce:ExecutionRoleArn":[
                "arn:aws:iam::111122223333:role/test-emr-demo4*"
            ]
```

```
        }
    }
}
```

Lanzar un clúster nuevo con un rol de tiempo de ejecución

Ahora que ya tiene los permisos necesarios, lance un nuevo clúster con un rol de tiempo de ejecución que pueda usar con su espacio de trabajo de Amazon EMR Studio.

Si ya ha lanzado un clúster nuevo con un rol de tiempo de ejecución, puede pasar directamente a la sección the section called “Utilizar el clúster con el espacio de trabajo”.

1. En primer lugar, complete los requisitos previos de la sección Roles en tiempo de ejecución para los pasos de Amazon EMR.
2. A continuación, lance un clúster con la siguiente configuración para usar los roles de tiempo de ejecución con los espacios de trabajo de Amazon EMR Studio. Para obtener instrucciones sobre cómo lanzar el clúster, consulte Especificar una configuración de seguridad para un clúster de Amazon EMR.
   - Seleccione la etiqueta de versión emr-6.11.0 o posterior.
   - Seleccione Spark, Livy y Jupyter Enterprise Gateway como aplicaciones de clúster.
   - Use la configuración de seguridad que creó en el paso anterior.
   - Si lo desea, puede habilitar Lake Formation para el clúster de EMR. Para obtener más información, consulte Habilitación de Lake Formation con Amazon EMR.

Después de lanzar el clúster, estará listo para usar el clúster habilitado para roles de tiempo de ejecución con un espacio de trabajo de EMR Studio.

> **Note**
>
> El ExecutionRoleArnvalor no es compatible actualmente con la operación de la StartNotebookExecutionAPI cuando el `ExecutionEngineConfig.Type` valor sí lo es`EMR`.

Utilizar el clúster de EMR con un rol de tiempo de ejecución en los espacios de trabajo

Después de configurar y lanzar el clúster, puede usar el clúster habilitado para roles de tiempo de ejecución con su espacio de trabajo de EMR Studio.

1. Cree un nuevo espacio de trabajo o lance uno existente. Para obtener más información, consulte Crear un espacio de trabajo de EMR Studio.
2. Elija la pestaña Clústeres de EMR en la barra lateral izquierda del espacio de trabajo abierto, expanda la sección Tipo de cómputo y elija su clúster en el menú del clúster de EMR y el rol de tiempo de ejecución en el EC2 menú del rol de tiempo de ejecución.
3. Seleccione Asociar para asociar el clúster con el rol de tiempo de ejecución a su espacio de trabajo.

> **Note**
>
> Cuando elija un rol en tiempo de ejecución, tenga en cuenta que puede tener políticas gestionadas subyacentes asociadas. En la mayoría de los casos, recomendamos elegir recursos limitados, como cuadernos específicos. Si, por ejemplo, eliges una función de tiempo de ejecución que incluya el acceso a todas tus libretas, la política gestionada asociada a la función proporciona acceso total.

Consideraciones

Tenga en cuenta las siguientes consideraciones cuando utilice un clúster habilitado para roles de tiempo de ejecución con su espacio de trabajo de Amazon EMR Studio:

- Solo puede seleccionar un rol de tiempo de ejecución al asociar un espacio de trabajo de EMR Studio a un clúster de EMR que utilice la versión 6.11 de Amazon EMR o superior.
- La funcionalidad de rol en tiempo de ejecución descrita en esta página solo es compatible con Amazon EMR que se ejecuta en Amazon EC2 y no es compatible con las aplicaciones interactivas EMR Serverless. Para obtener más información sobre los roles de tiempo de ejecución de EMR sin servidor, consulte Roles de tiempo de ejecución de tareas en la Guía del usuario de Amazon EMR sin servidor.
- Si bien necesita configurar permisos adicionales antes de poder especificar un rol de tiempo de ejecución al enviar un trabajo a un clúster, no necesita permisos adicionales para acceder a los archivos generados por un espacio de trabajo de EMR Studio. Los permisos de estos archivos son los mismos que los de los archivos generados a partir de los clústeres sin roles de tiempo de ejecución.

- No puede usar el Explorador de SQL en un espacio de trabajo de EMR Studio con un clúster que tenga un rol de tiempo de ejecución. Amazon EMR deshabilita el Explorador de SQL en la interfaz de usuario cuando un espacio de trabajo está conectado a un clúster de EMR habilitado para los roles de tiempo de ejecución.
- No puede usar el modo de colaboración en un espacio de trabajo de EMR Studio con un clúster que tenga un rol de tiempo de ejecución. Amazon EMR deshabilita las capacidades de colaboración en el espacio de trabajo cuando un espacio de trabajo está conectado a un clúster de EMR con roles en tiempo de ejecución. Solo podrá acceder al espacio de trabajo el usuario que lo asoció.
- No se pueden utilizar roles de tiempo de ejecución en un Studio con la propagación de identidades de confianza de IAM Identity Center habilitada.
- Es posible que aparezca la advertencia “¡Puede que la página no sea segura!” desde la interfaz de usuario de Spark para un clúster con funciones en tiempo de ejecución que utilice Amazon EMR versión 7.4.0 y versiones anteriores. En tal caso, omita la alerta para seguir viendo la interfaz de usuario de Spark.

Ejecutar los cuadernos de espacios de trabajo de Amazon EMR Studio mediante programación

> **Note**
>
> Las aplicaciones interactivas de Amazon EMR sin servidor no admiten la ejecución programática de cuadernos.

Puede ejecutar los cuadernos de sus espacios de trabajo de Amazon EMR Studio mediante programación con un script o en la AWS CLI. Para obtener información sobre cómo ejecutar su cuaderno mediante programación, consulte Ejemplos de comandos programáticos para Cuadernos de EMR.

Examinar los datos con SQL Explorer para EMR Studio

> **Note**
>
> SQL Explorer para EMR Studio no es compatible con las aplicaciones interactivas de Amazon EMR sin servidor o en un Studio que tenga habilitada la propagación de identidades de confianza de IAM Identity Center.

En este tema, se proporciona información que lo ayudará a comenzar a trabajar con el Explorador de SQL de Amazon EMR Studio. El Explorador de SQL es una herramienta de una sola página en su espacio de trabajo que le ayuda a comprender los orígenes de datos del catálogo de datos del clúster de EMR. Puede usar el Explorador de SQL para examinar sus datos, ejecutar consultas SQL para recuperar datos y descargar los resultados de las consultas.

El Explorador de SQL es compatible con Presto. Antes de usar el Explorador de SQL, asegúrese de que su clúster use la versión 5.34.0 o posterior o la versión 6.4.0 o posterior de Amazon EMR y tenga instalado Presto. El Explorador de SQL de Amazon EMR Studio no admite los clústeres de Presto que haya configurado con cifrado en tránsito. Esto se debe a que Presto se ejecuta en modo TLS en estos clústeres.

Explorar el catálogo de datos del clúster

El Explorador de SQL proporciona una interfaz de explorador del catálogo que puede utilizar para explorar y comprender cómo se organizan los datos. Por ejemplo, puede usar el explorador del catálogo de datos para comprobar los nombres de las tablas y columnas antes de escribir una consulta SQL.

Para examinar el catálogo de datos

1. Abra el Explorador de SQL en el espacio de trabajo.
2. Asegúrese de que su espacio de trabajo esté conectado a un clúster de EMR que se esté ejecutando y EC2 que utilice Amazon EMR versión 6.4.0 o posterior con Presto instalado. Puede elegir un clúster existente o crear uno nuevo. Para obtener más información, consulte Asociar computación a un espacio de trabajo de EMR Studio.
3. Seleccione una base de datos de la lista desplegable para explorarla.

4. Expanda una tabla de la base de datos para ver los nombres de las columnas de la tabla. También puede introducir una palabra clave en la barra de búsqueda para filtrar los resultados de la tabla.

Ejecutar una consulta SQL para recuperar datos

Para recuperar datos con una consulta SQL y descargar los resultados

1. Abra el Explorador de SQL en el espacio de trabajo.
2. Asegúrese de que su espacio de trabajo esté conectado a un clúster de EMR que se esté ejecutando EC2 con Presto y Spark instalados. Puede elegir un clúster existente o crear uno nuevo. Para obtener más información, consulte Asociar computación a un espacio de trabajo de EMR Studio.
3. Seleccione Abrir editor para abrir una nueva pestaña de editor en su espacio de trabajo.
4. Redacte la consulta SQL en la pestaña del editor.
5. Seleccione Ejecutar.
6. Consulte los resultados de la consulta en la vista previa de los resultados. El Explorador de SQL muestra los primeros 100 resultados de forma predeterminada. Puede elegir un número diferente de resultados para mostrarlos (hasta 1000) mediante el menú desplegable Vista previa de los primeros 100 resultados de la consulta.
7. Para descargar los resultados en formato CSV, seleccione Descargar resultados. Puede descargar hasta 1000 filas de resultados.

Asociar computación a un espacio de trabajo de EMR Studio

Amazon EMR Studio ejecuta comandos de cuadernos mediante un kernel en un clúster de EMR. Antes de poder seleccionar un núcleo, debe adjuntar el Workspace a un clúster que utilice EC2 instancias de Amazon, a un clúster de Amazon EMR en EKS o a una aplicación EMR Serverless. EMR Studio le permite asociar espacios de trabajo a clústeres nuevos o existentes y le brinda la flexibilidad de cambiar los clústeres sin cerrar el espacio de trabajo.

En esta sección se cubren los temas siguientes, que le ayudarán a trabajar con clústeres de EMR Studio y aprovisionarlos:

- Adjuntar un EC2 clúster de Amazon a un espacio de trabajo de EMR Studio
- Asociar un clúster de Amazon EMR en EKS a un espacio de trabajo de EMR Studio

- Asociar una aplicación de Amazon EMR sin servidor a un espacio de trabajo de EMR Studio
- Crear y asociar un nuevo clúster de EMR a un espacio de trabajo de EMR Studio
- Separar la computación de un espacio de trabajo de EMR Studio

Adjuntar un EC2 clúster de Amazon a un espacio de trabajo de EMR Studio

Puedes adjuntar un clúster de EMR que se ejecute en Amazon EC2 a un espacio de trabajo al crear el espacio de trabajo o adjuntar un clúster a un espacio de trabajo existente. Si desea crear y asociar un clúster nuevo, consulte Crear y asociar un nuevo clúster de EMR a un espacio de trabajo de EMR Studio.

> **Note**
>
> Un espacio de trabajo de Studio con la propagación de identidades de confianza de IAM Identity Center habilitada solo puede conectarse a un clúster de EMR con una configuración de seguridad que tenga Identity Center habilitado.

On create

Asociar un clúster de computación de Amazon EMR al crear un espacio de trabajo

1. En el cuadro de diálogo Crear un espacio de trabajo, asegúrese de haber seleccionado una subred para el nuevo espacio de trabajo. Amplíe la sección Configuración avanzada.
2. Seleccione Asociar espacio de trabajo a un clúster de EMR.
3. En la lista desplegable de clústeres de EMR, seleccione un clúster de EMR existente para asociarlo al espacio de trabajo.

Tras asociar un clúster, termine de crear el espacio de trabajo. Cuando abra el nuevo espacio de trabajo por primera vez y seleccione el panel Clústeres de EMR, debería ver el clúster seleccionado asociado.

On launch

Asociar un clúster de computación de Amazon EMR al iniciar el espacio de trabajo

1. Navegue hasta la lista de espacios de trabajo y seleccione la fila del espacio de trabajo que desee iniciar. A continuación, seleccione Iniciar espacio de trabajo > Iniciar con opciones.

2. Seleccione un clúster de EMR para asociarlo al espacio de trabajo.

Tras asociar un clúster, termine de crear el espacio de trabajo. Cuando abra el nuevo espacio de trabajo por primera vez y seleccione el panel Clústeres de EMR, debería ver el clúster seleccionado asociado.

In JupyterLab

Adjunte un espacio de trabajo a un clúster de procesamiento de Amazon EMR en JupyterLab

1. Seleccione su espacio de trabajo y, a continuación, seleccione Iniciar espacio de trabajo > Inicio rápido.
2. En el interior JupyterLab, abre la pestaña Clúster en la barra lateral izquierda.
3. Seleccione el menú desplegable EMR en el EC2 clúster o seleccione un clúster de Amazon EMR en EKS.
4. Seleccione Asociar para asociar el clúster al espacio de trabajo.

Tras asociar el clúster, termine de crear el espacio de trabajo. Cuando abra el nuevo espacio de trabajo por primera vez y seleccione el panel Clústeres de EMR, debería ver el clúster seleccionado asociado.

In the Workspace UI

Asocie un espacio de trabajo a un clúster de computación de Amazon EMR desde la interfaz de usuario del espacio de trabajo

1. En el espacio de trabajo que desee asociar a un clúster, elija el icono Clústeres de EMR en la barra lateral izquierda para abrir el panel Clúster.
2. En Tipo de clúster, expanda el menú desplegable y seleccione Clúster de EMR activado. EC2
3. Seleccione un clúster de la lista desplegable. Es posible que primero tenga que separar un clúster existente para habilitar la lista desplegable de selección del clúster.
4. Elija Adjuntar. Cuando el clúster esté asociado, debería aparecer un mensaje de confirmación.

Asociar un clúster de Amazon EMR en EKS a un espacio de trabajo de EMR Studio

Además de utilizar los clústeres de Amazon EMR que se ejecutan en Amazon EC2, puede adjuntar un Workspace a un clúster de Amazon EMR en EKS para ejecutar el código del bloc de notas. Para obtener más información sobre Amazon EMR en EKS, consulte Qué es Amazon EMR en EKS.

Para poder conectar un espacio de trabajo a un clúster de Amazon EMR en EKS, el administrador de Studio debe concederle permisos de acceso.

> **Note**
>
> No puede lanzar un clúster de Amazon EMR en EKS en un EMR Studio con la propagación de identidades de confianza de IAM Identity Center.

On create

Para asociar un clúster de Amazon EMR en EKS al crear un espacio de trabajo

1. En el cuadro de diálogo Crear un espacio de trabajo, amplíe la sección Configuración avanzada.
2. Seleccione Asociar espacio de trabajo a un clúster de Amazon EMR en EKS.
3. En Clúster de Amazon EMR en EKS, elija un clúster de la lista desplegable.
4. En Seleccione un punto de conexión, seleccione un punto de conexión administrado para asociarlo al espacio de trabajo. Un punto de conexión administrado es una puerta de enlace que permite a EMR Studio comunicarse con el clúster elegido.
5. Seleccione Crear un espacio de trabajo para finalizar el proceso de creación del espacio de trabajo y asociar el clúster seleccionado.

Tras asociar un clúster, puede terminar el proceso de creación del espacio de trabajo. Cuando abra el nuevo espacio de trabajo por primera vez y seleccione el panel Clústeres de EMR, debería ver que el clúster seleccionado está asociado.

In the Workspace UI

Para asociar un espacio de trabajo a un clúster de Amazon EMR en EKS desde la interfaz de usuario del espacio de trabajo

1. En el espacio de trabajo que desee asociar a un clúster, elija el icono Clústeres de EMR en la barra lateral izquierda para abrir el panel Clúster.
2. Expanda el menú desplegable Tipo de clúster y seleccione Clústeres de EMR en EKS.
3. En Clúster de EMR en EKS, seleccione un clúster de la lista desplegable.
4. En Punto de conexión, seleccione un punto de conexión administrado para asociarlo al espacio de trabajo. Un punto de conexión administrado es una puerta de enlace que permite a EMR Studio comunicarse con el clúster elegido.
5. Elija Adjuntar. Cuando el clúster esté asociado, debería aparecer un mensaje de confirmación.

Asociar una aplicación de Amazon EMR sin servidor a un espacio de trabajo de EMR Studio

Puede asociar un espacio de trabajo a una aplicación de EMR sin servidor para ejecutar cargas de trabajo interactivas. Para obtener más información, consulte Uso de cuadernos para ejecutar cargas de trabajo interactivas con EMR sin servidor a través de EMR Studio.

> **Note**
>
> No puede conectar una aplicación de EMR sin servidor a un EMR Studio con la propagación de identidades de confianza de IAM Identity Center.

Example Adjunte un espacio de trabajo a una aplicación EMR sin servidor en JupyterLab

Para que pueda conectar un espacio de trabajo a una aplicación de EMR sin servidor, el administrador de su cuenta debe concederle los permisos de acceso, tal y como se describe en Required permissions for interactive workloads.

1. Diríjase a EMR Studio, seleccione su espacio de trabajo y, a continuación, seleccione Iniciar espacio de trabajo > Inicio rápido.
2. En el interior JupyterLab, abre la pestaña Clúster en la barra lateral izquierda.

3. Seleccione EMR sin servidor como opción de computación y, a continuación, seleccione una aplicación de EMR sin servidor y un rol de tiempo de ejecución.
4. Seleccione Asociar para asociar el clúster al espacio de trabajo.

Ahora, cuando abra este espacio de trabajo, debería ver la aplicación seleccionada que asoció.

Crear y asociar un nuevo clúster de EMR a un espacio de trabajo de EMR Studio

Los usuarios avanzados de EMR Studio pueden aprovisionar nuevos clústeres de EMR que se ejecuten en Amazon EC2 para usarlos con un espacio de trabajo. El nuevo clúster tiene todas las aplicaciones de macrodatos necesarias para EMR Studio instaladas de forma predeterminada.

Para crear clústeres, el administrador de Studio primero debe concederle el permiso mediante una política de sesión. Para obtener más información, consulte Creación de políticas de permisos para los usuarios de EMR Studio.

Puede crear un clúster nuevo en el cuadro de diálogo Crear un espacio de trabajo o desde el panel Clúster de la interfaz de usuario del espacio de trabajo. En cualquier caso, tiene dos opciones para crear un clúster:

1. Crear un clúster de EMR: cree un clúster de EMR eligiendo el tipo y el recuento de EC2 instancias de Amazon.
2. Utilizar una plantilla de clúster: aprovisione un clúster seleccionando una plantilla de clúster predefinida. Esta opción aparece si tiene permiso para usar plantillas de clúster.

> **Note**
>
> Si la propagación de identidades de confianza de IAM Identity Center está habilitada en el Studio, debe utilizar una plantilla para crear un clúster.

Para crear un clúster de EMR proporcionando una configuración de clúster

1. Elija un punto de inicio.

Para…	Haga lo siguiente...
Cree el clúster al crear un espacio de trabajo con el cuadro de diálogo Crear un espacio de trabajo.	Amplíe la sección Configuración avanzada en el cuadro de diálogo Crear un espacio de trabajo y seleccione Crear un clúster de EMR.
Cree el clúster desde el panel de clústeres de EMR en la interfaz de usuario del espacio de trabajo después de haber creado un espacio de trabajo.	Elija la pestaña Clústeres de EMR en la barra lateral izquierda de un espacio de trabajo abierto, expanda la sección Configura ción avanzada y elija Crear clúster.

2. Ingrese un nombre de clúster. Asignar un nombre al clúster le ayudará a encontrarlo más adelante en la lista de clústeres de EMR Studio.
3. En Versión de Amazon EMR, seleccione una versión de Amazon EMR para el clúster.
4. Por ejemplo, selecciona el tipo y el número de EC2 instancias de Amazon para el clúster. Para obtener más información sobre la selección de los tipos de instancia, consulte Configurar los tipos de EC2 instancias de Amazon para utilizarlos con Amazon EMR. Se utilizará una instancia como nodo principal.
5. Seleccione una subred en la que EMR Studio pueda lanzar el nuevo clúster. El administrador de Studio aprueba previamente cada opción de subred. Su espacio de trabajo debería poder conectarse a un clúster de cualquier subred de la lista.
6. Seleccione un URI de S3 para el almacenamiento de registros.
7. Seleccione Crear clúster de EMR para aprovisionarlo. Si usa el cuadro de diálogo Crear un espacio de trabajo, seleccione Crear un espacio de trabajo para crear el espacio de trabajo y aprovisionar el clúster. Una vez que EMR Studio aprovisione el nuevo clúster, lo asocia al espacio de trabajo.

Para crear un clúster mediante una plantilla de clúster

1. Elija un punto de inicio.

Para…	Haga lo siguiente...
Cree el clúster al crear un espacio de trabajo con el cuadro de diálogo Crear un espacio de trabajo.	Expanda la sección Configuración avanzada en el cuadro de diálogo Crear un espacio de trabajo y seleccione Usar una plantilla de clúster.
Cree el clúster desde el panel de clústeres de EMR en la interfaz de usuario del espacio de trabajo.	Seleccione la pestaña Clústeres de EMR en la barra lateral izquierda de un espacio de trabajo abierto, expanda la sección Configura ción avanzada y seleccione Plantilla de clúster.

2. En la lista desplegable, seleccione una plantilla de clúster. Cada plantilla de clúster disponible incluye una breve descripción para ayudarle a realizar una selección.
3. La plantilla de clúster que elija puede tener parámetros adicionales, como la versión de lanzamiento de Amazon EMR o el nombre del clúster. Puede elegir o insertar valores, o bien utilizar los valores predeterminados que haya seleccionado el administrador.
4. Seleccione una subred en la que EMR Studio pueda lanzar el nuevo clúster. El administrador de Studio aprueba previamente cada opción de subred. Su espacio de trabajo debería poder conectarse a un clúster de cualquier subred.
5. Seleccione Usar plantilla de clúster para aprovisionar el clúster y asociarlo al espacio de trabajo. EMR Studio tardará unos minutos en crear el clúster. Si usa el cuadro de diálogo Crear un espacio de trabajo, seleccione Crear un espacio de trabajo para crear el espacio de trabajo y aprovisionar el clúster. Una vez que EMR Studio aprovisione el nuevo clúster, lo asocia al espacio de trabajo.

Separar la computación de un espacio de trabajo de EMR Studio

Para intercambiar el clúster asociado a un espacio de trabajo, puede separar un clúster de la interfaz de usuario del espacio de trabajo.

Para separar un clúster de un espacio de trabajo

1. En el espacio de trabajo que desee separar de un clúster, elija el icono Clústeres de EMR en la barra lateral izquierda para abrir el panel Clúster.

2. En Seleccionar clúster, seleccione Separar y espere a que EMR Studio separe el clúster. Cuando se separe el clúster, verá un mensaje indicándole que la operación se ha realizado correctamente.

Para separar una aplicación de EMR sin servidor de un espacio de trabajo de EMR Studio

Para intercambiar la computación asociada a un espacio de trabajo, puede separar la aplicación de la interfaz de usuario del espacio de trabajo.

1. En el espacio de trabajo que desee separar de un clúster, elija el icono Computación de Amazon EMR en la barra lateral izquierda para abrir el panel Computación.
2. En Seleccionar computación, seleccione Separar y espere a que EMR Studio separe la aplicación. Cuando se separe la aplicación, verá un mensaje indicándole que la operación se ha realizado correctamente.

Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio

Asocie hasta tres repositorios basados en Git a un espacio de trabajo de Amazon EMR Studio para guardar y compartir archivos de cuaderno.

Acerca de los repositorios de Git para EMR Studio

Puede asociar un máximo de tres repositorios de Git a un espacio de trabajo de EMR Studio. De forma predeterminada, cada espacio de trabajo te permite elegir de una lista de repositorios de Git que están asociados a la misma AWS cuenta que Studio. También puede crear un nuevo repositorio de Git como recurso para un espacio de trabajo.

Puede ejecutar comandos de Git como los siguientes mediante un comando de terminal mientras está conectado al nodo principal de un clúster.

```
!git pull origin <branch-name>
```

También puede utilizar la extensión jupyterlab-git. Ábralo desde la barra lateral izquierda seleccionando el icono de Git. Para obtener información sobre la extensión jupyterlab-git, consulta jupyterlab-git. JupyterLab

Requisitos previos

- Para asociar un repositorio de Git a un espacio de trabajo, el estudio debe estar configurado para permitir la vinculación de repositorios de Git. El administrador de Studio debe tomar medidas para Establecer el acceso y los permisos para los repositorios basados en Git.
- Si utilizas un CodeCommit repositorio, debes usar las credenciales de Git y HTTPS. No se admiten las claves SSH y HTTPS con el asistente de AWS Command Line Interface credenciales. CodeCommit tampoco admite los tokens de acceso personal (). PATs Para obtener más información, consulte Uso de IAM con CodeCommit en la Guía del usuario de IAM y Configuración para usuarios de HTTPS que utilizan credenciales de Git en la Guía del AWS CodeCommit usuario.

Instrucciones

Para vincular un repositorio de Git asociado a un espacio de trabajo

1. Abra el espacio de trabajo que quiera vincular a un repositorio desde la lista Espacios de trabajo del estudio.
2. En la barra lateral izquierda, seleccione el icono del repositorio de Git de Amazon EMR para abrir el panel de herramientas del repositorio de Git.
3. En Repositorios de Git, expanda la lista desplegable y seleccione un máximo de tres repositorios para vincularlos al espacio de trabajo. EMR Studio registra su selección y comienza a vincular cada repositorio.

Es posible que el proceso de vinculación tarde algún tiempo en completarse. Puede ver el estado de cada repositorio que haya seleccionado en el panel de herramientas del repositorio de Git. Después de que EMR Studio vincule un repositorio a un espacio de trabajo, debería ver los archivos que pertenecen a ese repositorio en el panel del Explorador de archivos.

Para agregar un nuevo repositorio de Git a un espacio de trabajo como recurso

1. Abra el espacio de trabajo que quiera vincular a un repositorio desde la lista de espacios de trabajo del estudio.
2. En la barra lateral izquierda, seleccione el icono del repositorio de Git de Amazon EMR para abrir el panel de herramientas del repositorio de Git.
3. Seleccione Agregar nuevo repositorio de Git.

4. En Nombre del repositorio, escriba un nombre descriptivo para el repositorio de EMR Studio. Los nombres pueden contener caracteres alfanuméricos, guiones o guiones bajos.
5. En Git repository URL (URL del repositorio de Git), escriba la URL del repositorio. Cuando utilizas un CodeCommit repositorio, esta es la URL que se copia cuando eliges Clonar URL y, a continuación, clonar HTTPS. Por ejemplo, `https://git-codecommit.us-west-2.amazonaws.com/v1/repos/[MyCodeCommitRepoName]`.
6. En Ramificación, introduzca el nombre de una ramificación existente que quiera seleccionar.
7. En Credenciales de Git, seleccione una opción de acuerdo con las siguientes pautas. EMR Studio accede a sus credenciales de Git mediante los secretos almacenados en Secrets Manager.

> **Note**
>
> Si utilizas un GitHub repositorio, te recomendamos que utilices un token de acceso personal (PAT) para autenticarte. A partir del 13 de agosto de 2021, GitHub se requerirá la autenticación basada en token y ya no se aceptarán contraseñas al autenticar las operaciones de Git. Para obtener más información, consulta la publicación sobre los requisitos de autenticación de token para las operaciones de Git en El GitHub Blog.

Opción	Descripción
Crear un nuevo secreto	Elige esta opción para asociar las credencia les de Git existentes a un nuevo secreto que se AWS Secrets Manager creará automátic amente. Realice una de las siguientes acciones basadas en las credenciales de Git que utilice para el repositorio. Si usa un nombre de usuario y una contraseñ a de Git para acceder al repositorio, seleccione Nombre de usuario y contraseñ a, escriba el nombre del secreto que va a usar en Secrets Manager y, a continuación, escriba el nombre de usuario y la contraseña que asociar al secreto.

Opción	Descripción
	–O BIEN– Si utiliza un token de acceso personal para acceder al repositorio, seleccione Token de acceso personal (PAT), escriba el nombre del secreto que usará en Secrets Manager y, a continuación, escriba su token de acceso personal. Para obtener más información, consulta Crear un token de acceso personal para la línea de comandos GitHub y Tokens de acceso personal para Bitbucket. CodeCommit Los repositorios no admiten esta opción.
Usar un repositorio público sin credenciales	Elija esta opción para acceder a un repositorio público.
Usa un secreto existente AWS	Elija esta opción si ya ha guardado sus credenciales como un secreto en Secrets Manager y, a continuación, seleccione el nombre del secreto en la lista. Si selecciona un secreto asociado con un nombre de usuario y contraseña de Git, el secreto debe estar en el formato `{"gitUsername": " MyUserName ", "gitPassword": " MyPassword "}`.

8. Seleccione Agregar repositorio para crear el nuevo repositorio. Después de que EMR Studio cree el nuevo repositorio, verá un mensaje de confirmación. El nuevo repositorio aparece en la lista desplegable de Repositorios de Git.
9. Para vincular el nuevo repositorio a su espacio de trabajo, selecciónelo en la lista desplegable de Repositorios de Git.

Es posible que el proceso de vinculación tarde algún tiempo en completarse. Después de que EMR Studio vincule el nuevo repositorio al espacio de trabajo, debería aparecer una nueva carpeta con el mismo nombre que el repositorio en el panel del Explorador de archivos.

Para abrir un repositorio vinculado diferente, navegue hasta su carpeta en el Explorador de archivos.

Uso del editor SQL de Amazon Athena en EMR Studio

Descripción general

Puede usar Amazon EMR Studio para desarrollar y ejecutar consultas interactivas en Amazon Athena. Esto significa que puede realizar análisis de SQL en Athena desde la misma interfaz de EMR Studio que utiliza para ejecutar sus cargas de trabajo de Spark, Scala y otras. Con esta integración, puedes usar la función de autocompletado para desarrollar consultas rápidamente, buscar datos en tu catálogo de datos de AWS Glue, crear consultas guardadas, ver tu historial de consultas y mucho más.

Para obtener más información sobre cómo utilizar Amazon Athena, consulte Uso de Athena SQL en la Guía del usuario de Amazon Athena.

Uso del editor SQL de Athena en EMR Studio

Realice los siguientes pasos para desarrollar y ejecutar consultas interactivas en Amazon Athena desde su EMR Studio:

1. Agregue los permisos necesarios al rol de usuario para los usuarios que accedan a los espacios de trabajo de este Studio. Los permisos se muestran en la tabla AWS Identity and Access Management permisos para los usuarios de EMR Studio en la columna Acceda al editor SQL de Amazon Athena desde EMR Studio. Como alternativa, puede optar por copiar el contenido de la política Avanzada de Ejemplo de políticas de usuario para conceder a los usuarios permisos completos a las capacidades de EMR Studio, incluida esta.
2. Configure y cree un EMR Studio.
3. En el Studio, seleccione el Editor de consultas en la barra lateral.

Debería poder ver la conocida interfaz de usuario del editor Athena. Para obtener información sobre cómo empezar a utilizar Athena SQL para ejecutar consultas interactivas, consulte Introducción y Uso de Athena SQL en la Guía del usuario de Amazon Athena.

> **Note**
>
> Si la propagación de identidades de confianza de IAM Identity Center está habilitada en el EMR Studio, debe utilizar los grupos de trabajo de Athena para controlar el acceso a las consultas y el grupo de trabajo que utilice también debe utilizar la propagación de identidades de confianza. Para ver los pasos de configuración de Identity Center y habilitación de la propagación de identidades de confianza para su grupo de trabajo, consulte Uso de los grupos de trabajo de Athena habilitados para IAM Identity Center en la Guía del usuario de Amazon Athena.

Consideraciones sobre el uso del editor SQL de Athena en EMR Studio

- La integración con Athena está disponible en todos las regiones comerciales donde están disponibles EMR Studio y Athena.
- Las siguientes características de Athena no están disponibles en EMR Studio:
  - Características administrativas, por ejemplo, crear o actualizar grupos de trabajo, orígenes de datos o reservas de capacidad de Athena
  - Athena para Spark o portátiles Spark
  - DataZone Integración con Amazon
  - Optimizador basado en costes (CBO)
  - Step Functions

CodeWhisperer Integración de Amazon con EMR Studio Workspaces

Descripción general

Puede utilizar Amazon CodeWhisperer con Amazon EMR Studio para obtener recomendaciones en tiempo real a medida que escribe código. JupyterLab CodeWhispererpuede completar sus comentarios, terminar líneas de código individuales, hacer line-by-line recomendaciones y generar funciones con formato completo.

> **Note**
>
> Cuando utiliza Amazon EMR Studio, es AWS posible que almacene datos sobre su uso y contenido con el fin de mejorar el servicio. Para obtener más información e instrucciones

para optar por no compartir datos, consulta Compartir tus datos con AWS en la Guía del CodeWhisperer usuario de Amazon.

Consideraciones para su uso CodeWhisperer con Workspaces

- CodeWhisperer la integración está disponible en el mismo Regiones de AWS lugar donde está disponible EMR Studio, como se documenta en las consideraciones de EMR Studio.
- Amazon EMR Studio utiliza automáticamente el CodeWhisperer punto de conexión en EE. UU. Este (Norte de Virginia) (us-east-1) para las recomendaciones, independientemente de la región en la que se encuentre su estudio.
- CodeWhisperer solo admite el lenguaje Python para codificar scripts ETL para trabajos de Spark en EMR Studio.
- Una opción de telemetría del lado del cliente cuantifica el uso de. CodeWhisperer No se admite esta funcionalidad con EMR Studio.

Se requieren permisos para CodeWhisperer

Para utilizarla CodeWhisperer, debe adjuntar la siguiente política a su rol de usuario de IAM para Amazon EMR Studio:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "CodeWhispererPermissions",
      "Effect": "Allow",
      "Action": [
        "codewhisperer:GenerateRecommendations"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Úselo CodeWhisperer con espacios de trabajo

Para mostrar el registro de CodeWhisperer referencia JupyterLab, abra el CodeWhispererpanel en la parte inferior de la JupyterLab ventana y elija Abrir registro de referencia de código.

La siguiente lista contiene atajos que puede utilizar para interactuar con las CodeWhisperer sugerencias:

- Recomendaciones de pausa: usa la opción Pausar las sugerencias automáticas desde la CodeWhisperer configuración.
- Aceptar una recomendación: presione la tecla Tab en el teclado.
- Rechazar una recomendación: presione la tecla Escape en el teclado.
- Navegar por las recomendaciones: utilice las flechas arriba y abajo del teclado.
- Invocar manualmente: presione las teclas Alt y C en el teclado. Si utiliza un dispositivo Mac, presione Cmd y C.

También puedes usarlo CodeWhisperer para cambiar ajustes como el nivel de registro y obtener sugerencias de referencias de código. Para obtener más información, consulta Configuración CodeWhisperer JupyterLab y funciones en la Guía del CodeWhisperer usuario de Amazon.

Depurar aplicaciones y trabajos con EMR Studio

Con Amazon EMR Studio, puede lanzar interfaces de aplicaciones de datos para analizar las aplicaciones y las ejecuciones de trabajos en el navegador.

También puede lanzar las interfaces de usuario persistentes y fuera del clúster para Amazon EMR que se ejecutan EC2 en clústeres desde la consola de Amazon EMR. Para obtener más información, consulte Visualización de interfaces de usuario de aplicaciones persistentes en Amazon EMR.

> **Note**
>
> En función de la configuración del navegador, es posible que necesite habilitar las ventanas emergentes para que se abra la interfaz de usuario de una aplicación.

Para obtener información sobre la configuración y el uso de las interfaces de la aplicación, consulte The YARN Timeline Server, Monitoring and Instrumentation o Tez UI Overview.

Depure Amazon EMR que se ejecuta en Amazon Jobs EC2

Workspace UI

Iniciar una interfaz de usuario en el clúster a partir de un archivo de cuaderno

Si utiliza las versiones 5.33.0 y posteriores de Amazon EMR, puede iniciar la interfaz de usuario web de Spark (la interfaz de usuario de Spark o el servidor de historial de Spark) desde un cuaderno en el espacio de trabajo.

UIs Trabaje en el clúster con los PySpark núcleos Spark o SparkR. El tamaño máximo de archivo visible para los registros de eventos o contenedores de Spark es de 10 MB. Si sus archivos de registro superan los 10 MB, le recomendamos que utilice el servidor de historial de Spark persistente en lugar de la interfaz de usuario de Spark integrada en el clúster para depurar los trabajos.

> **Important**
>
> Para que EMR Studio pueda lanzar interfaces de usuario de aplicaciones en un clúster desde un espacio de trabajo, un clúster debe poder comunicarse con Amazon API Gateway. Debe configurar el clúster de EMR para permitir el tráfico de red saliente a Amazon API Gateway y asegurarse de que se pueda acceder a Amazon API Gateway desde el clúster.
> La interfaz de usuario de Spark accede a los registros del contenedor resolviendo los nombres de host. Si utiliza un nombre de dominio personalizado, debe asegurarse de que Amazon DNS o el servidor DNS que especifique puedan resolver los nombres de host de los nodos de su clúster. Para ello, defina las opciones del protocolo de configuración dinámica de host (DHCP) para la Amazon Virtual Private Cloud (VPC) asociada al clúster. Para obtener más información sobre las opciones de DHCP, consulte los conjuntos de opciones de DHCP en la Guía del usuario de Amazon Virtual Private Cloud.

1. En su EMR Studio, abra el espacio de trabajo que desee usar y asegúrese de que está conectado a un clúster de Amazon EMR en ejecución. EC2 Para obtener instrucciones, consulte Asociar computación a un espacio de trabajo de EMR Studio.
2. Abre un archivo de bloc de notas y usa el PySpark núcleo Spark o SparkR. Para seleccionar un kernel, seleccione su nombre en la parte superior derecha de la barra de herramientas del

cuaderno para abrir el cuadro de diálogo Seleccionar kernel. El nombre aparece como No hay ningún kernel si no se ha seleccionado ningún kernel.

3. Ejecute el código de su cuaderno. El fragmento siguiente aparece como salida en el cuaderno al iniciar el contexto de Spark. Puede que tarde unos segundos en aparecer. Si ha iniciado el contexto de Spark, puede ejecutar el comando `%%info` para acceder a un enlace que le llevará a la interfaz de usuario de Spark en cualquier momento.

> **Note**
>
> Si los enlaces de la interfaz de usuario de Spark no funcionan o no aparecen después de unos segundos, cree una nueva celda de cuaderno y ejecute el comando `%%info` para regenerar los enlaces.

4. Para iniciar la interfaz de usuario de Spark, seleccione Vincular en la interfaz de usuario de Spark. Si la aplicación de Spark se está ejecutando, la interfaz de usuario de Spark se abrirá en una pestaña nueva. Si la aplicación se ha completado, en su lugar se abrirá el servidor de historial de Spark.

   Tras iniciar la interfaz de usuario de Spark, puedes modificar la URL en el navegador para abrir el YARN ResourceManager o el servidor de cronología de Yarn. Agregue una de las siguientes rutas después de `amazonaws.com`.

Interfaz de usuario web	Ruta	Ejemplo de URL modificada
YARN ResourceManager	/rm	*j-examplebby5ij* https://.emrappui-prod. *eu-west-1* .amazonaws.com */rm*
Yarn Timeline Server	/yts	https://.emrappui-prod. *j-examplebby5ij* *eu-west-1* .amazonaws.com */yts*
Servidor de historial de Spark	/shs	https://.emrappui-prod. *j-examplebby5ij* *eu-west-1* .amazonaws.com */shs*

Studio UI

Iniciar el servidor YARN Timeline, el servidor de historial de Spark o la interfaz de usuario de Tez persistentes desde la interfaz de usuario de EMR Studio

1. En su EMR Studio, seleccione Amazon EMR en la parte izquierda de la EC2 página para abrir la lista Amazon EMR en clústeres. EC2
2. Filtre la lista de clústeres por nombre, estado o identificador introduciendo valores en el cuadro de búsqueda. También puede buscar por intervalo de tiempo de creación.
3. Seleccione un clúster y, a continuación, elija Lanzar aplicación UIs para seleccionar una interfaz de usuario de la aplicación. La interfaz de usuario de la aplicación se abre en una nueva pestaña del navegador y puede que tarde en cargarse.

Depurar un EMR Studio en ejecución en EMR sin servidor

Al igual que Amazon EMR que se ejecuta en Amazon EC2, puede utilizar la interfaz de usuario de Workspace para analizar sus aplicaciones EMR sin servidor. En la interfaz de usuario del espacio de trabajo, si utiliza las versiones 6.14.0 y posteriores de Amazon EMR, puede iniciar la interfaz de usuario web de Spark (la interfaz de usuario de Spark o el servidor de historial de Spark) desde un cuaderno en el espacio de trabajo. Para su comodidad, también ofrecemos un enlace para acceder rápidamente a los registros de controladores de Spark.

Depurar Amazon EMR en ejecuciones de trabajos de EKS con el servidor de historial de Spark

Cuando envía una ejecución de trabajo a un clúster de Amazon EMR en EKS, puede acceder a los registros de esa ejecución de trabajos mediante el servidor de historial de Spark. El servidor de historial de Spark proporciona herramientas para monitorear las aplicaciones de Spark, como una lista de las etapas y tareas del programador, un resumen del tamaño de los RDD y el uso de memoria, e información sobre el entorno. Puede iniciar el servidor de historial de Spark para las ejecuciones de trabajos de Amazon EMR en EKS de las siguientes maneras:

- Cuando envíe una ejecución de trabajo con EMR Studio con un punto de conexión administrado de Amazon EMR en EKS, puede lanzar el servidor de historial de Spark desde un archivo de cuaderno de su espacio de trabajo.
- Cuando envíes una ejecución de trabajo con el AWS SDK AWS CLI o Amazon EMR en EKS, puedes lanzar el Spark History Server desde la interfaz de usuario de EMR Studio.

Para obtener información sobre cómo utilizar el servidor de historial de Spark, consulte Supervisión e instrumentación en la documentación de Apache Spark. Para obtener más información sobre las ejecuciones de trabajos, consulte Conceptos y componentes en la Guía de desarrollo de Amazon EMR en EKS.

Para iniciar el servidor de historial de Spark desde un archivo de cuaderno en su espacio de trabajo de EMR Studio

1. Abra un espacio de trabajo que esté conectado a un clúster de Amazon EMR en EKS.
2. Seleccione y abra el archivo de su cuaderno en el espacio de trabajo.
3. Elija la interfaz de usuario de Spark en la parte superior del archivo del cuaderno para abrir el servidor de historial de Spark persistente en una pestaña nueva.

Para iniciar el servidor de historial de Spark desde la interfaz de usuario de EMR Studio

> **Note**
>
> La lista de trabajos de la interfaz de usuario de EMR Studio muestra solo las ejecuciones de trabajos que envíe mediante el uso de Amazon EMR AWS CLI o el AWS SDK para Amazon EMR en EKS.

1. En su EMR Studio, seleccione Amazon EMR en EKS en la parte izquierda de la página.
2. Busque el clúster virtual de Amazon EMR en EKS que utilizó para enviar la ejecución de su trabajo. Puede filtrar la lista de clústeres por estado o identificador introduciendo valores en el cuadro de búsqueda.
3. Seleccione el clúster para abrir su página de detalles. En la página de detalles se muestra información sobre el clúster, como el identificador, el espacio de nombres y el estado. En la página también se muestra una lista de todas las ejecuciones de trabajos enviadas a ese clúster.
4. En la página de detalles del clúster, seleccione una ejecución de trabajos para depurarla.
5. En la parte superior derecha de la lista Trabajos, seleccione Iniciar servidor de historial de Spark para abrir la interfaz de la aplicación en una nueva pestaña del navegador.

Instalar kernels y bibliotecas en un espacio de trabajo de EMR Studio

Cada espacio de trabajo de Amazon EMR Studio viene con un conjunto de bibliotecas y kernels preinstalados.

Núcleos y bibliotecas en clústeres que se ejecutan en Amazon EC2

También puede personalizar el entorno de EMR Studio de las siguientes maneras cuando utilice clústeres de EMR que se ejecutan en Amazon: EC2

- Instale los kernels de cuadernos de Jupyter y las bibliotecas de Python en el nodo principal de un clúster: al instalar bibliotecas con esta opción, todos los espacios de trabajo asociados al mismo clúster comparten esas bibliotecas. Puede instalar kernels o bibliotecas desde una celda del cuaderno o mientras está conectado mediante SSH al nodo principal de un clúster.
- Usar bibliotecas para cuadernos: cuando los usuarios del espacio de trabajo instalan y usan bibliotecas desde la celda de un cuaderno, esas bibliotecas solo están disponibles para ese cuaderno. Esta opción permite que diferentes cuadernos que utilizan el mismo clúster funcionen sin tener que preocuparse por los conflictos en las versiones de las bibliotecas.

Los espacios de trabajo de EMR Studio tienen la misma arquitectura subyacente que Cuadernos de EMR. Puede instalar y utilizar los kernels de cuadernos de Jupyter y las bibliotecas de Python con EMR Studio del mismo modo que lo haría con Cuadernos de EMR. Para obtener instrucciones, consulte Instalación y uso de kernels y bibliotecas en EMR Studio.

Kernels y bibliotecas de los clústeres de Amazon EMR en EKS

Los clústeres de Amazon EMR en EKS incluyen los núcleos PySpark Python 3.7 y Python 3.7 con un conjunto de bibliotecas preinstaladas. Amazon EMR en EKS no admite la instalación de bibliotecas o clústeres adicionales.

Cada clúster de Amazon EMR en EKS viene con las siguientes PySpark bibliotecas y Python instaladas:

- Python – boto3, cffi, future, ggplot, jupyter, kubernetes, matplotlib, numpy, pandas, plotly, pycryptodomex, py4j, requests, scikit-learn, scipy, seaborn
- PySpark – ggplot, jupyter, matplotlib, numpy, pandas, plotly, pycryptodomex, py4j, requests, scikit-learn, scipy, seaborn

Kernels y bibliotecas en aplicaciones de EMR sin servidor

Cada aplicación EMR Serverless viene con las siguientes bibliotecas y PySpark Python instaladas:

- Python – ggplot, matplotlib, numpy, pandas, plotly, bokeh, scikit-learn, scipy, seaborn
- PySpark – ggplot, matplotlib,numpy, pandas, plotly, bokeh, scikit-learn, scipy, seaborn

Mejorar los kernels con comandos magic en EMR Studio

Descripción general

EMR Studio y EMR Notebooks admiten comandos. magic Magiclos comandos, o magic comandos, son mejoras que el IPython núcleo proporciona para ayudarle a ejecutar y analizar datos. IPythones un entorno de shell interactivo creado con Python.

Amazon EMR también admite Sparkmagic un paquete que proporciona núcleos relacionados con Spark (PySparknúcleos SparkR y Scala) con comandos magic específicos y que usa Livy en el clúster para enviar trabajos de Spark.

Puede usar comandos magic siempre que tenga un kernel de Python en su cuaderno de EMR. Del mismo modo, cualquier kernel relacionado con Spark admite los comandos de Sparkmagic.

Los comandos Magic, también llamados magics, vienen en dos variedades:

- magics de línea: estos comandos magic se indican con un único `%` como prefijo y funcionan en una sola línea de código.
- magics de celda: estos comandos magic se indican con un doble `%%` como prefijo y funcionan en varias líneas de código.

Para ver todos los comandos magic disponibles, consulte Enumerar los comandos magic y Sparkmagic.

Consideraciones y limitaciones

- EMR sin servidor no admite `%%sh` para ejecutar `spark-submit`. No es compatible con magic de Cuadernos de EMR.
- Los clústeres de Amazon EMR en EKS no admiten los comandos de Sparkmagic para EMR Studio. Esto se debe a que los kernels de Spark que se utilizan con los puntos de conexión administrados están integrados en Kubernetes, y Sparkmagic y Livy no los admiten. Como

solución alternativa, puedes establecer la configuración de Spark directamente en el SparkContext objeto, como se muestra en el siguiente ejemplo.

```
spark.conf.set("spark.driver.maxResultSize", '6g')
```

- Los siguientes magic comandos y acciones están prohibidos por AWS:
  - `%alias`
  - `%alias_magic`
  - `%automagic`
  - `%macro`
  - Modificar `proxy_user` con `%configure`
  - Modificar `KERNEL_USERNAME` con `%env` o `%set_env`

Enumerar los comandos magic y Sparkmagic

Utilice los siguientes comandos para enumerar los comandos magic disponibles:

- `%lsmagic` muestra una lista de todas las funciones magic disponibles actualmente.
- `%%help` muestra las funciones magic relacionadas con Spark que están disponibles actualmente en el paquete de Sparkmagic.

Usar `%%configure` para configurar Spark

Uno de los comandos más útiles de Sparkmagic es el comando `%%configure`, que configura los parámetros de creación de sesiones. En cuanto a los ajustes de `conf`, puede establecer cualquier configuración de Spark que se mencione en la documentación de configuración de Apache Spark.

Example Agregar un archivo JAR externo a los Cuadernos de EMR desde el repositorio de Maven o Amazon S3

Puede usar el siguiente enfoque para agregar una dependencia de un archivo JAR externo a cualquier kernel relacionado con Spark compatible con Sparkmagic.

```
%%configure -f
{"conf": {
    "spark.jars.packages": "com.jsuereth:scala-arm_2.11:2.0,ml.combust.bundle:bundle-
ml_2.11:0.13.0,com.databricks:dbutils-api_2.11:0.0.3",
```

```
    "spark.jars": "s3://amzn-s3-demo-bucket/my-jar.jar"
    }
}
```

Example : Configurar Hudi

Puede utilizar el editor de cuadernos para configurar el cuaderno de EMR de modo que utilice Hudi.

```
%%configure
{ "conf": {
     "spark.jars": "hdfs://apps/hudi/lib/hudi-spark-bundle.jar,hdfs:///apps/hudi/lib/
spark-spark-avro.jar",
     "spark.serializer": "org.apache.spark.serializer.KryoSerializer",
     "spark.sql.hive.convertMetastoreParquet":"false"
     }
}
```

Usar `%%sh` para ejecutar `spark-submit`

La magic `%%sh` ejecuta comandos del intérprete de comandos en un subproceso en una instancia del clúster asociado. Normalmente, utilizaría uno de los kernels relacionados con Spark para ejecutar las aplicaciones de Spark en su clúster asociado. Sin embargo, si quiere usar un kernel de Python para enviar una aplicación de Spark, puede usar la siguiente magic y reemplazar el nombre del bucket por el nombre de su bucket en minúsculas.

```
%%sh
spark-submit --master yarn --deploy-mode cluster s3://amzn-s3-demo-bucket/test.py
```

En este ejemplo, el clúster tiene que acceder a la ubicación de `s3://amzn-s3-demo-bucket/test.py` o se producirá un error en el comando.

Con la magic `%%sh`, puede usar cualquier comando de Linux. Si desea ejecutar algún comando de Spark o YARN, use una de las siguientes opciones para crear un usuario de Hadoop de `emr-notebook` y concederle permisos para ejecutar los comandos:

- Para crear un usuario nuevo de forma explícita, ejecute los siguientes comandos.

  ```
  hadoop fs -mkdir /user/emr-notebook
  hadoop fs -chown emr-notebook /user/emr-notebook
  ```

- Puede activar la suplantación de usuarios en Livy, lo que crea el usuario automáticamente. Para obtener más información, consulte Habilitar la suplantación de usuario para supervisar la actividad del usuario y del trabajo de Spark.

Usar `%%display` para visualizar los marcos de datos de Spark

Puede usar magic `%%display` para ver un marco de datos de Spark. Para usar esta magic, ejecute el siguiente comando.

```
%%display df
```

Elija ver los resultados en formato de tabla, como se muestra en la siguiente imagen.

También puede optar por visualizar sus datos con cinco tipos de gráficos. Las opciones son gráficos circulares, de dispersión, de líneas, de áreas y de barras.

Usar la magic de Cuadernos de EMR

Amazon EMR proporciona la siguiente magic de Cuadernos de EMR que puede utilizar con los kernels basados en Python3 y Spark:

- `%mount_workspace_dir`: monta el directorio del espacio de trabajo en el clúster para que pueda importar y ejecutar código desde otros archivos de su espacio de trabajo

  > **Note**
  >
  > Con `%mount_workspace_dir`, solo el kernel de Python 3 puede acceder a sus sistemas de archivos locales. Los ejecutores de Spark no tendrán acceso al directorio montado con este kernel.

- `%umount_workspace_dir`: desmonta el directorio del espacio de trabajo del clúster
- `%generate_s3_download_url`: genera un enlace de descarga temporal en la salida del cuaderno para un objeto de Amazon S3

Requisitos previos

Antes de instalar la magic de Cuadernos de EMR, complete las siguientes tareas:

- Asegúrese de que Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia) tenga acceso de lectura a Amazon S3. `EMR_EC2_DefaultRole` con la política administrada `AmazonElasticMapReduceforEC2Role` cumple con este requisito. Si usa un rol o una política personalizados, asegúrese de que tenga los permisos de S3 necesarios.

> **Note**
>
> Los EMR magic Notebooks se ejecutan en un clúster como usuario del bloc de notas y utilizan EC2 el perfil de instancia para interactuar con Amazon S3. Al montar un directorio del espacio de trabajo en un clúster de EMR, todos los espacios de trabajo y cuadernos de EMR con permiso para asociarse a ese clúster pueden acceder al directorio montado. De forma predeterminada, los directorios se montan como de solo lectura. Si bien `s3fs-fuse` y `goofys` permiten los montajes de lectura y escritura, le recomendamos especialmente que no modifique los parámetros de montaje para montar directorios en modo de lectura y escritura. Si permite el acceso de escritura, todos los cambios que se realicen en el directorio se escribirán en el bucket de S3. Para evitar que se eliminen o sobrescriban accidentalmente, puede habilitar el control de versiones en su bucket de S3. Para obtener más información, consulte Uso del control de versiones en buckets de S3.

- Ejecute uno de los siguientes scripts en su clúster para instalar las dependencias de la magic de Cuadernos de EMR. Para ejecutar un script, puede Usar acciones de arranque personalizadas o seguir las instrucciones de Ejecutar comandos y scripts en un clúster de Amazon EMR cuando ya tenga un clúster en ejecución.

  Puede elegir qué dependencia desea instalar. Tanto s3fs-fuse como goofys son herramientas FUSE (sistema de archivos en el espacio de usuario) que permiten montar un bucket de Amazon S3 como un sistema de archivos local en un clúster. La herramienta `s3fs` proporciona una experiencia similar a POSIX. La herramienta `goofys` es una buena opción si prefiere el rendimiento en lugar de un sistema de archivos compatible con POSIX.

  La serie Amazon EMR 7.x usa Amazon Linux 2023, que no es compatible con los repositorios EPEL. Si utiliza Amazon EMR 7.x, siga las instrucciones de s3fs-fuse para realizar la instalación GitHub. `s3fs-fuse` Si usa las series 5.x o 6.x, utilice los siguientes comandos para instalar `s3fs-fuse`.

```
#!/bin/sh

# Install the s3fs dependency for EMR Notebooks magics
```

```
sudo amazon-linux-extras install epel -y
sudo yum install s3fs-fuse -y
```

OR

```
#!/bin/sh

# Install the goofys dependency for EMR Notebooks magics
sudo wget https://github.com/kahing/goofys/releases/latest/download/goofys -P /usr/
bin/
sudo chmod ugo+x /usr/bin/goofys
```

Instalar la magic de Cuadernos de EMR

> **Note**
>
> Con las versiones de 6.0 a 6.9.0 y de 5.0 a 5.36.0 de Amazon EMR, solo las versiones 0.2.0 y superiores del paquete `emr-notebooks-magics` admiten la magic `%mount_workspace_dir`.

Siga los pasos que se indican a continuación para instalar la magic de Cuadernos de EMR.

1. En su cuaderno, ejecute los siguientes comandos para instalar el paquete `emr-notebooks-magics`.

   ```
   %pip install boto3 --upgrade
   %pip install botocore --upgrade
   %pip install emr-notebooks-magics --upgrade
   ```

2. Reinicie el kernel para cargar la magic de Cuadernos de EMR.
3. Verifique la instalación con el siguiente comando, que debería mostrar el texto de ayuda de la salida de `%mount_workspace_dir`.

   ```
   %mount_workspace_dir?
   ```

Montar un directorio del espacio de trabajo con **`%mount_workspace_dir`**

La magic `%mount_workspace_dir` le permite montar el directorio del espacio de trabajo en su clúster de EMR para que pueda importar y ejecutar otros archivos, módulos o paquetes almacenados en su directorio.

En el siguiente ejemplo, se monta todo el directorio del espacio de trabajo en un clúster y se especifica el argumento opcional *`<--fuse-type>`* para usar goofys para montar el directorio.

```
%mount_workspace_dir . <--fuse-type goofys>
```

Para comprobar que el directorio del espacio de trabajo esté montado, utilice el siguiente ejemplo para mostrar el directorio de trabajo actual con el comando `ls`. La salida debe mostrar todos los archivos del espacio de trabajo.

```
%%sh
ls
```

Cuando termine de realizar cambios en su espacio de trabajo, puede desmontar el directorio del espacio de trabajo con el siguiente comando:

> **Note**
>
> El directorio del espacio de trabajo permanece montado en el clúster incluso cuando el espacio de trabajo está detenido o separado. Debe desmontar el directorio del espacio de trabajo de forma explícita.

```
%umount_workspace_dir
```

Descargar un objeto de Amazon S3 con **`%generate_s3_download_url`**

El comando `generate_s3_download_url`, crea una URL prefirmada para un objeto almacenado en Amazon S3. Puede usar la URL prefirmada para descargar el objeto a su máquina local. Por ejemplo, puede ejecutar `generate_s3_download_url` para descargar el resultado de una consulta SQL que su código escribe en Amazon S3.

De forma predeterminada, la URL prefirmada es válida durante 60 minutos. Puede cambiar el tiempo de caducidad especificando un número de segundos para la marca `--expires-in`. Por ejemplo, `--expires-in 1800` crea una URL que es válida durante 30 minutos.

El siguiente ejemplo genera un enlace de descarga para un objeto especificando la ruta completa de Amazon S3: *`s3://EXAMPLE-DOC-BUCKET/path/to/my/object`*.

```
%generate_s3_download_url s3://EXAMPLE-DOC-BUCKET/path/to/my/object
```

Para obtener más información sobre el uso de `generate_s3_download_url`, ejecute el siguiente comando para mostrar el texto de ayuda.

```
%generate_s3_download_url?
```

Ejecución de un cuaderno en el modo sin encabezado con **`%execute_notebook`**

Con la magic `%execute_notebook`, puede ejecutar otro cuaderno en el modo sin encabezado y ver el resultado de cada celda que haya ejecutado. Esto magic requiere permisos adicionales para el rol de instancia que Amazon EMR y Amazon EC2 comparten. Para obtener más información sobre cómo conceder permisos adicionales, ejecute el comando `%execute_notebook?`.

Durante un trabajo de larga duración, es posible que el sistema entre en modo de suspensión debido a la inactividad o que pierda temporalmente la conectividad a internet. Esto podría interrumpir la conexión entre su navegador y el servidor de Jupyter. En este caso, podría perder la salida de las celdas que ha ejecutado y enviado desde el servidor de Jupyter.

Si ejecuta el cuaderno en el modo sin encabezado con la magic `%execute_notebook`, Cuadernos de EMR captura el resultado de las celdas que se han ejecutado, incluso si la red local sufre una interrupción. Cuadernos de EMR guarda la salida de forma incremental en un nuevo cuaderno con el mismo nombre que el cuaderno que ha utilizado. A continuación, Cuadernos de EMR coloca el cuaderno en una nueva carpeta dentro del espacio de trabajo. Las ejecuciones sin encabezado se realizan en el mismo clúster y utilizan el rol de servicio `EMR_Notebook_DefaultRole`, pero los argumentos adicionales pueden alterar los valores predeterminados.

Para ejecutar un cuaderno en el modo sin encabezado, utilice el siguiente comando:

```
%execute_notebook <relative-file-path>
```

Para especificar un ID de clúster y un rol de servicio para una ejecución sin encabezado, use el siguiente comando:

```
%execute_notebook <notebook_name>.ipynb --cluster-id <emr-cluster-id> --service-role
 <emr-notebook-service-role>
```

Cuando Amazon EMR y Amazon EC2 comparten un rol de instancia, el rol requiere los siguientes permisos adicionales:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:StartNotebookExecution",
        "elasticmapreduce:DescribeNotebookExecution",
        "ec2:DescribeInstances"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowELASTICMAPREDUCEStartnotebookexecution"
    },
    {
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/EMR_Notebooks_DefaultRole"
      ],
      "Sid": "AllowIAMPassrole"
    }
  ]
}
```

> **Note**
>
> Para usar magic `%execute_notebook`, instale la versión 0.2.3 o una superior del paquete `emr-notebooks-magics`.

Usar cuadernos multilingües con kernels de Spark

Cada kernel de cuadernos de Jupyter tiene un lenguaje predeterminado. Por ejemplo, el idioma predeterminado del núcleo de Spark es Scala y el idioma predeterminado de PySpark los núcleos es Python. Con Amazon EMR 6.4.0 y las versiones posteriores, EMR Studio admite los cuadernos de varios lenguajes. Esto significa que cada kernel de EMR Studio puede admitir los siguientes lenguajes además del lenguaje predeterminado: Python, Spark, R y Spark SQL.

Para activar esta característica, especifique uno de los siguientes comandos magic al principio de cualquier celda.

Idioma	Comando
Python	`%%pyspark`
Scala	`%%scalaspark`
R	`%%rspark` No se admite para cargas de trabajo interactivas con EMR sin servidor.
Spark SQL	`%%sql`

Cuando se invocan, estos comandos ejecutan toda la celda dentro de la misma sesión de Spark utilizando el intérprete del lenguaje correspondiente.

La `%%pyspark` celda magic permite a los usuarios escribir PySpark código en todos los núcleos de Spark.

```
%%pyspark
a = 1
```

La magic de celda `%%sql` permite a los usuarios ejecutar código Spark-SQL en todos los kernels de Spark.

```
%%sql
SHOW TABLES
```

La magic de celda `%%rspark` permite a los usuarios ejecutar código SparkR en todos los kernels de Spark.

```
%%rspark
a <- 1
```

La magic de celda `%%scalaspark` permite a los usuarios ejecutar código Spark Scala en todos los kernels de Spark.

```
%%scalaspark
val a = 1
```

Compartir datos entre intérpretes de lenguajes con tablas temporales

También puede compartir datos entre intérpretes de lenguajes mediante tablas temporales. En el siguiente ejemplo, se usa `%%pyspark` en una celda para crear una tabla temporal en Python y se usa `%%scalaspark` en la celda siguiente para leer los datos de esa tabla en Scala.

```
%%pyspark
df=spark.sql("SELECT * from nyc_top_trips_report LIMIT 20")
# create a temporary table called nyc_top_trips_report_view in python
df.createOrReplaceTempView("nyc_top_trips_report_view")
```

```
%%scalaspark
// read the temp table in scala
val df=spark.sql("SELECT * from nyc_top_trips_report_view")
df.show(5)
```

Información general de Cuadernos de Amazon EMR

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Puede utilizar Amazon EMR Notebooks junto con clústeres de Amazon EMR que ejecuten Apache Spark para crear y abrir Jupyter Notebook e JupyterLab interfaces dentro de la consola de Amazon EMR. Un cuaderno de EMR es un cuaderno “sin servidor” que puede usar para ejecutar consultas y código. A diferencia de un cuaderno tradicional, el contenido de un cuaderno de EMR (ecuaciones, consultas, modelos, código y texto narrativo dentro de las celdas del cuaderno) se ejecutan en un cliente. Los comandos se ejecutan utilizando un kernel en el clúster de EMR. El contenido del cuaderno también se guarda en Amazon S3 separado de los datos del clúster para mayor durabilidad y flexibilidad en la reutilización.

Puede iniciar un clúster, asociar un cuaderno de EMR para análisis y, a continuación, terminar el clúster. También puede cerrar un bloc de notas asociado a un clúster en ejecución y cambiar a otro. Varios usuarios pueden asociar cuadernos al mismo clúster de forma simultánea y compartir entre sí archivos de cuaderno en Amazon S3. Estas características le permiten ejecutar clústeres bajo demanda para ahorrar costes y reducir el tiempo dedicado a reconfigurar blocs de notas para diferentes clústeres y conjuntos de datos.

También puede ejecutar un cuaderno de EMR mediante programación con la API de Amazon EMR, sin necesidad de interactuar con la consola de Amazon EMR (“ejecución Headless”). Debe incluir una celda en el cuaderno de EMR que tenga una etiqueta de parámetros. Esa celda permite que un script pase nuevos valores de entrada al cuaderno. Los cuadernos parametrizados se pueden reutilizar con diferentes conjuntos de valores de entrada. No es necesario hacer copias del mismo cuaderno para editarlo y ejecutarlo con nuevos valores de entrada. Amazon EMR crea y guarda el cuaderno de salida en S3 para cada ejecución del cuaderno parametrizado. Para ver muestras de códigos de la API de cuadernos de EMR, consulte Ejemplos de comandos programáticos para Cuadernos de EMR.

> **Important**
>
> La capacidad de Cuadernos de EMR admite clústeres que utilizan la versión 5.18.0 de Amazon EMR y versiones posteriores. Le recomendamos que utilice Cuadernos de EMR con clústeres que utilicen la última versión de Amazon EMR, o al menos las versiones 5.30.0, 5.32.0 o 6.2.0. Con estas versiones, los kernels de Jupyter se ejecutan en el clúster asociado, en lugar de hacerlo en una instancia de Jupyter. Este cambio mejora el rendimiento y mejora su capacidad para personalizar kernels y bibliotecas. Para obtener más información, consulte Diferencias en capacidades por versión de clúster.

Se aplicarán los cargos correspondientes para el almacenamiento de Amazon S3 y los clústeres de Amazon EMR.

Cuadernos de Amazon EMR está disponible como Espacios de trabajo de Amazon EMR Studio en la consola

Hacer la transición de Cuadernos de EMR a Espacios de trabajo

En la nueva consola de Amazon EMR, hemos fusionado Cuadernos de EMR con Espacios de trabajo de Amazon EMR Studio en una sola experiencia. Cuando utiliza un EMR Studio, puede crear y configurar diferentes espacios de trabajo para organizar y ejecutar cuadernos. Si tenía cuadernos de Amazon EMR en la consola anterior, estos están disponibles como espacios de trabajo de EMR Studio en la consola.

Amazon EMR creó estos nuevos espacios de trabajo de EMR Studio. El número de estudios que hemos creado corresponde al número de distintos VPCs que utiliza en los Notebooks EMR. Por ejemplo, si se conecta a clústeres de EMR en dos Notebooks diferentes VPCs a los de EMR, crearemos dos nuevos EMR Studios. Sus cuadernos se distribuyen entre los nuevos Studios.

> **Important**
>
> Desactivamos la opción para crear nuevos cuadernos en la antigua consola de Amazon EMR. En su lugar, utilice Crear espacio de trabajo en la nueva consola de Amazon EMR.

Para obtener más información sobre Espacios de trabajo de Amazon EMR Studio, consulte Conozca los espacios de trabajo de EMR Studio. Para obtener información general conceptual sobre EMR Studio, consulte Workspaces en la página Cómo funciona Amazon EMR Studio.

¿Qué necesita hacer?

Si bien puede seguir utilizando sus cuadernos actuales en la consola anterior, le recomendamos que utilice Espacios de trabajo de Amazon EMR Studio en la consola. Debe configurar permisos de rol adicionales para activar las capacidades en EMR Studio que no están disponibles en Cuadernos de EMR.

> **Note**
>
> Como mínimo, para ver los cuadernos de EMR existentes como espacios de trabajo de EMR Studio y para crear nuevos espacios de trabajo, los usuarios deben tener los permisos `elasticmapreduce:ListStudios` y `elasticmapreduce:CreateStudioPresignedUrl` en sus roles. Para acceder a todas las características de EMR Studio, consulte Habilitación de las características de EMR Studio para los usuarios de Cuadernos de EMR para ver la lista completa de permisos adicionales que necesitarán los usuarios de Cuadernos de EMR.

Capacidades mejoradas en EMR Studio más allá de los cuadernos de EMR

Con Amazon EMR Studio, puede configurar y utilizar las siguientes capacidades que no están disponibles en Cuadernos de EMR:

- Explorar clústeres de EMR y adjuntar a ellos desde JupyterLab
- Explorar clústeres virtuales de Cuadernos de EMR y adjuntar a ellos desde JupyterLab
- Conectarse a repositorios de Git desde JupyterLab
- Colaborar con otros miembros de su equipo para escribir y ejecutar código de cuaderno
- Examinar los datos con SQL Explorer
- Aprovisionar clústeres de EMR con Service Catalog

Para obtener una lista completa de las capacidades de Amazon EMR Studio, consulte Características principales de EMR Studio.

Habilitación de las características de EMR Studio para los usuarios de Cuadernos de EMR

Los nuevos EMR Studios que crearemos como parte de esta fusión utilizan el rol de IAM `EMR_Notebooks_DefaultRole` existente como rol de servicio de EMR Studio.

Los usuarios que lleven a cabo la transición a EMR Studio desde Cuadernos de EMR y deseen utilizar las capacidades adicionales de EMR Studio necesitarán varios permisos de rol nuevos. Agregue los siguientes permisos a los roles de los usuarios de Cuadernos de EMR que planean usar EMR Studio.

> **Note**
>
> Como mínimo, para ver los cuadernos de EMR existentes como espacios de trabajo de EMR Studio y para crear nuevos espacios de trabajo, los usuarios deben tener los permisos `elasticmapreduce:ListStudios` y `elasticmapreduce:CreateStudioPresignedUrl` en sus roles. Para utilizar todas las características de EMR Studio, agregue todos los permisos que se indican a continuación. Los usuarios administradores también necesitan permiso para crear y administrar un EMR Studio. Para obtener más información, consulte Permisos de administrador para crear y administrar un EMR Studio.

```
"elasticmapreduce:DescribeStudio",
"elasticmapreduce:ListStudios",
"elasticmapreduce:CreateStudioPresignedUrl",
"elasticmapreduce:UpdateEditor",
"elasticmapreduce:PutWorkspaceAccess",
"elasticmapreduce:DeleteWorkspaceAccess",
"elasticmapreduce:ListWorkspaceAccessIdentities",
"emr-containers:ListVirtualClusters",
"emr-containers:DescribeVirtualCluster",
"emr-containers:ListManagedEndpoints",
"emr-containers:DescribeManagedEndpoint",
"emr-containers:CreateAccessTokenForManagedEndpoint",
"emr-containers:ListJobRuns",
"emr-containers:DescribeJobRun",
"servicecatalog:SearchProducts",
"servicecatalog:DescribeProduct",
"servicecatalog:DescribeProductView",
```

```
"servicecatalog:DescribeProvisioningParameters",
"servicecatalog:ProvisionProduct",
"servicecatalog:UpdateProvisionedProduct",
"servicecatalog:ListProvisioningArtifacts",
"servicecatalog:DescribeRecord",
"servicecatalog:ListLaunchPaths",
"cloudformation:DescribeStackResources"
```

Los siguientes permisos también son necesarios para usar las capacidades de colaboración de EMR Studio, pero no eran necesarios con Cuadernos de EMR.

```
"sso-directory:SearchUsers",
"iam:GetUser",
"iam:GetRole",
"iam:ListUsers",
"iam:ListRoles",
"sso:GetManagedApplicationInstance"
```

Requisitos, diferencias en las versiones de lanzamiento y seguridad de los cuadernos de EMR

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Tenga en cuenta los siguientes requisitos, las diferencias en las versiones de las versiones, la información de seguridad y otras consideraciones al crear clústeres y desarrollar soluciones mediante los cuadernos de EMR.

Requisitos del clúster

- Habilite el acceso público por bloques de Amazon EMR: el acceso entrante a un clúster permite a los usuarios del clúster ejecutar kernels de cuadernos. Asegúrese de que solo los usuarios

autorizados puedan acceder al clúster. Recomendamos encarecidamente que deje habilitado el acceso público de bloqueo y que limite el tráfico SSH entrante solo a fuentes de confianza. Para obtener más información, consulte Uso de Bloquear el acceso público de Amazon EMR y Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

- Utilice un clúster compatible: un clúster conectado a un cuaderno debe cumplir los siguientes requisitos:
  - Solo se admiten clústeres creados mediante Amazon EMR. Puede crear un clúster de forma independiente en Amazon EMR y, a continuación, asociar un cuaderno de EMR, o puede crear un clúster compatible al crear un cuaderno de EMR.
  - Solo se admiten los clústeres creados con la versión 5.18.0 de Amazon EMR o versiones posteriores. Consulte the section called “Diferencias en capacidades por versión de clúster”.
  - No se admiten los clústeres creados con EC2 instancias de Amazon con procesadores AMD EPYC (por ejemplo, los tipos de instancia m5a.* y r5a.*).
  - Cuadernos de EMR solo funciona con clústeres creados con `VisibleToAllUsers` establecidos en `true`. `VisibleToAllUsers` es `true` de forma predeterminada.
  - El clúster debe lanzarse dentro de una EC2 -VPC. Se admiten subredes públicas y privadas. La plataforma EC2 -Classic no es compatible.
  - El clúster se debe lanzar con Hadoop, Spark y Livy instalados. Se pueden instalar otras aplicaciones, pero Cuadernos de EMR actualmente solo admite clústeres de Spark.

    > **Important**
    >
    > Para las versiones 5.32.0 y posteriores de Amazon EMR, o 6.2.0 y posteriores, su clúster también debe ejecutar la aplicación Jupyter Enterprise Gateway para poder funcionar con Cuadernos de EMR.

  - No se admiten los clústeres que utilizan la autenticación de Kerberos.
  - Los clústeres integrados AWS Lake Formation solo admiten la instalación de bibliotecas para ordenadores portátiles. En el clúster no se admite la instalación de kernels ni bibliotecas.
  - No se admiten clústeres con varios nodos principales.
  - No se admiten los clústeres que utilizan EC2 instancias de Amazon basadas en AWS Graviton2.

Diferencias en capacidades por versión de clúster

Le recomendamos que utilice Cuadernos de EMR con clústeres creados con las versiones 5.30.0, 5.32.0 o posteriores, o 6.2.0 o posteriores de Amazon EMR. Con estas versiones, Cuadernos de EMR ejecuta los kernels en el clúster Amazon EMR adjunto. Los kernels y las bibliotecas se pueden instalar directamente en el nodo principal del clúster. El uso de Cuadernos de EMR con estas versiones de clúster tiene las siguientes ventajas:

- Rendimiento mejorado: los núcleos de los equipos portátiles se ejecutan en clústeres con los tipos de EC2 instancias que tú selecciones. Las versiones anteriores ejecutan kernels en una instancia especializada a la que no se puede cambiar el tamaño, acceder o personalizar.
- Posibilidad de agregar y personalizar kernels: puede conectarse al clúster para instalar los paquetes del kernel mediante `conda` y `pip`. Además, se admite la instalación de `pip` mediante comandos de terminal dentro de celdas de bloc de notas. En versiones anteriores, solo estaban disponibles los núcleos preinstalados (Python PySpark, Spark y SparkR). Para obtener más información, consulte Instalación de kernels y bibliotecas de Python en un nodo principal del clúster.
- Capacidad para instalar bibliotecas de Python: puede instalar bibliotecas de Python en el nodo principal del clúster mediante `conda` y `pip`. Recomendamos utilizar `conda`. En las versiones anteriores, solo se admitían las bibliotecas para ordenadores portátiles. PySpark

Funciones de Cuadernos de EMR admitidas por la versión del clúster

Versión de lanzamiento del clúster	Bibliotecas con formato de cuaderno para PySpark	Instalación del kernel en el clúster	Instalación de la biblioteca de Python en el nodo principal
Antes de 5.18.0	Cuadernos de EMR no es compatible		
5.18.0–5.25.0	No	No	No
5.26.0–5.29.0	Sí	No	No
5.30.0	Sí	Sí	Sí
6.0.0	No	No	No

Versión de lanzamiento del clúster	Bibliotecas con formato de cuaderno para PySpark	Instalación del kernel en el clúster	Instalación de la biblioteca de Python en el nodo principal
5.32.0 y versiones posteriores, y 6.2.0 y versiones posteriores	Sí	Sí	Sí

Límites para cuadernos asociados de forma simultánea

Al crear un clúster compatible con cuadernos, tenga en cuenta el tipo de EC2 instancia del nodo principal del clúster. Las limitaciones de memoria de esta EC2 instancia determinan el número de cuadernos que pueden estar preparados simultáneamente para ejecutar código y consultas en el clúster.

Tipo de EC2 instancia de nodo principal	Número de Cuadernos de EMR
*.medium	2
*.large	4
*.xlarge	8
*.2xlarge	16
*.4xlarge	24
*.8xlarge	24
*.16xlarge	24

Versiones de cuaderno de Jupyter y Python

Cuadernos de EMR ejecuta la versión 6.0.2 de Cuaderno de Jupyter y Python 3.6.5 independientemente de la versión de Amazon EMR del clúster asociado.

Consideraciones en torno a la seguridad

Uso de ubicaciones de S3 cifradas

Si especifica una ubicación cifrada en Amazon S3 para almacenar archivos de cuaderno, debe configurar Rol de servicio para Cuadernos de Amazon EMR como usuario clave. El rol de servicio predeterminado es `EMR_Notebooks_DefaultRole`. Si usa una AWS KMS clave para el cifrado, consulte Uso de políticas de claves en AWS KMS en la Guía para AWS Key Management Service desarrolladores y el artículo de soporte para agregar usuarios clave.

Uso de cookies con dominios de alojamiento

Para aumentar la seguridad de las aplicaciones fuera de la consola que podría utilizar con Amazon EMR, los dominios de alojamiento de aplicaciones se registran en la lista de sufijos públicos (PSL). Algunos ejemplos de estos dominios de alojamiento son los siguientes: `emrstudio-prod.us-east-1.amazonaws.com`, `emrnotebooks-prod.us-east-1.amazonaws.com`, `emrappui-prod.us-east-1.amazonaws.com`. Para mayor seguridad, si necesita configurar cookies confidenciales en el nombre de dominio predeterminado, le recomendamos que utilice cookies con el prefijo `__Host-`. Esta práctica lo ayuda a proteger su dominio de los intentos de falsificación de solicitudes entre sitios (CSRF). Para obtener más información, consulte la página Set-Cookie en la Red de desarrolladores de Mozilla.

Crear un cuaderno en EMR Studio

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Puede crear un cuaderno de EMR con la antigua consola de Amazon EMR. No se admite la AWS CLI creación de blocs de notas mediante la API Amazon EMR.

Para crear un bloc de notas de EMR

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com/elasticmapreduce/.
2. Elija Notebooks (Blocs de notas), Create notebook (Crear bloc de notas).
3. Introduzca el Notebook name (Nombre del bloc de notas) y una Notebook description (Descripción del bloc de notas) opcional.
4. Si tiene un clúster activo al que desea asociar el cuaderno, deje el valor predeterminado Elegir un clúster existente seleccionado, haga clic en Elegir, seleccione un clúster en la lista y, a continuación, haga clic en Elegir clúster. Para obtener información sobre los requisitos de clúster para Cuadernos de EMR, consulte Requisitos, diferencias en las versiones de lanzamiento y seguridad de los cuadernos de EMR.

   —o bien—

   Elija Create a cluster (Crear un clúster), introduzca un Cluster name (Nombre de clúster) y elija opciones según las siguientes directrices. El clúster se crea en la VPC predeterminada para la cuenta mediante instancias bajo demanda.

Opción	Descripción
Cluster name (Nombre del clúster)	Nombre descriptivo utilizado para identificar el clúster.
Release (Versión)	No se puede modificar. El valor predeterm inado es la versión de Amazon EMR más reciente (5.36.2).
Aplicaciones	No se puede modificar. Enumera las aplicaciones que están instaladas en el clúster.
Instancia	Introduzca el número de instancias y seleccione el tipo de EC2 instancia. Una instancia se utiliza para el nodo principal. El resto se utiliza para nodos principales. El tipo de instancia determina el número de blocs de notas que se pueden asociar simultáneamente al clúster. Para obtener

Opción	Descripción
	más información, consulte Límites para cuadernos asociados de forma simultánea.
Rol de EMR	Deje el valor predeterminado o elija el vínculo para el que desea especificar un rol de servicio personalizado de Amazon EMR. Para obtener más información, consulte Rol de servicio para Amazon EMR (rol de EMR).
EC2 perfil de instancia	Deje el valor predeterminado o elija el enlace para especificar un rol de servicio personali zado para EC2 las instancias. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia).
EC2 par de claves	Elija un par de EC2 claves para poder conectarse a las instancias del clúster. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.
Terminación automática	La terminación automática es compatible con las versiones 5.30.0 y 6.1.0 y posteriores de Amazon EMR. Seleccione la casilla de verificación para habilitar la terminación automática y, a continuación, especifique la cantidad de tiempo de inactividad tras el cual el clúster debe apagarse automáticamente. Para obtener más información, consulte Uso de una política de finalización automática para la limpieza de clústeres de Amazon EMR..

5. En Security groups (Grupos de seguridad), elija Use default security groups (Usar grupos de seguridad predeterminados). También puede elegir Choose security groups (Elegir grupos de

seguridad) y seleccionar grupos de seguridad personalizados disponibles en la VPC del clúster. Seleccione uno para la instancia principal y otro para la instancia de cliente de cuaderno. Para obtener más información, consulte the section called “Grupos de seguridad para Cuadernos de Amazon EMR”.

6. En Rol de servicio de AWS , deje el valor predeterminado o elija un rol personalizado en la lista. La instancia de cliente para el bloc de notas utiliza este rol. Para obtener más información, consulte Rol de servicio para Cuadernos de Amazon EMR.
7. En Ubicación del cuaderno, elija la ubicación de Amazon S3 donde se guarda el archivo del cuaderno o especifique su propia ubicación. Si el bucket y la carpeta no existen, Amazon EMR los creará.

   Amazon EMR crea una carpeta cuyo nombre es el valor indicado en ID del cuaderno y guarda el cuaderno en un archivo denominado `NotebookName.ipynb`. Por ejemplo, si especifica la ubicación de Amazon S3 `s3://amzn-s3-demo-bucket/MyNotebooks` para un cuaderno denominado `MyFirstEMRManagedNotebook`, el archivo del cuaderno se guardará en `s3://amzn-s3-demo-bucket/MyNotebooks/NotebookID/MyFirstEMRManagedNotebook.ipynb`.

   Si especifica una ubicación cifrada en Amazon S3, debe configurar Rol de servicio para Cuadernos de Amazon EMR como usuario de claves. El rol de servicio predeterminado es `EMR_Notebooks_DefaultRole`. Si usa una AWS KMS clave para el cifrado, consulte Uso de políticas de claves en AWS KMS en la Guía para AWS Key Management Service desarrolladores y el artículo de soporte para agregar usuarios clave.
8. Opcionalmente, si ha agregado un repositorio basado en Git a Amazon EMR que desea asociar con este cuaderno, elija Repositorio de Git, haga clic en Elegir repositorio y, a continuación, seleccione un repositorio de la lista. Para obtener más información, consulte Asociación de repositorios basados en Git con Cuadernos de EMR.
9. También puede elegir Tags (Etiquetas) y, a continuación, añadir etiquetas de clave-valor adicionales para el bloc de notas.

> **Important**
>
> Se aplicará una etiqueta predeterminada con la cadena Key (Clave) establecida en `creatorUserID` y el valor definido como su ID de usuario de IAM para poder obtener acceso. Recomendamos que no cambie ni elimine esta etiqueta, ya que se puede utilizar

para controlar el acceso. Para obtener más información, consulte Uso de etiquetas de clúster y cuaderno con políticas de IAM para el control de acceso.

10. Elija Create Notebook (Crear bloc de notas).

Uso de los Cuadernos de EMR

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Después de crear un cuaderno de EMR, el cuaderno tarda unos momentos en iniciarse. El campo Status (Estado) de la lista Notebooks (Blocs de notas) muestra Starting (Comenzando). Puede abrir un bloc de notas cuando su estado sea Ready (Listo). Es posible que el bloc de notas tarde algo más de tiempo en estar Ready (Listo) si lo ha creado al mismo tiempo que un clúster.

> **Tip**
>
> Actualice el navegador o elija el icono de actualización situado encima de la lista de blocs de notas para actualizar el estado del bloc de notas.

Descripción de los estados del cuaderno

Un cuaderno de EMR puede tener los siguientes valores en el campo Estado de la lista Cuadernos.

Estado	Significado
Ready	Puede abrir el bloc de notas con el editor de blocs de notas. Mientras el bloc de notas tenga el estado Ready (Listo), puede detenerlo o

Estado	Significado
	eliminarlo. Para cambiar de clúster, primero debe detener el bloc de notas. Si un bloc de notas que se encuentra en el estado Ready (Listo) está inactivo durante un periodo prolongado de tiempo, se detiene automátic amente.
Inicio	El bloc de notas se está creando y asociando al clúster. Mientras se inicia un bloc de notas, no se puede abrir el editor de blocs de notas, detenerlo, eliminarlo ni cambiar de clúster.
Pendiente	El bloc de notas se ha creado y está esperando a que finalice la integración con el clúster. El clúster puede estar todavía aprovisionando recursos o respondiendo a otras solicitudes. Puede abrir el editor de blocs de notas con el bloc de notas en modo local. El código que se basa en los procesos del clúster no se ejecuta y genera un error.
Detención	El bloc de notas se está cerrando o el clúster al que está asociado se está terminando. Mientras se detiene un bloc de notas, no se puede abrir el editor de blocs de notas, detenerlo, eliminarlo ni cambiar de clúster.
Stopped	El bloc de notas se ha cerrado. Puede iniciar el bloc de notas en el mismo clúster, siempre que este siga ejecutándose. Puede cambiar de clúster y eliminar el clúster.

Estado	Significado
Eliminación	El clúster se está eliminando de la lista de clústeres disponibles. El archivo del cuaderno, `NotebookName .ipynb` , permanece en Amazon S3 y continúa acumulando los cargos de almacenamiento correspondientes.

Uso del editor de cuadernos

Una ventaja de usar un cuaderno EMR es que puedes iniciar el cuaderno en Jupyter o JupyterLab directamente desde la consola.

Con EMR Notebooks, el editor de cuadernos al que puede acceder desde la consola Amazon EMR es el conocido editor Jupyter Notebook de código abierto o. JupyterLab Dado que el editor de cuadernos se lanza en la consola de Amazon EMR, resulta más eficaz configurar el acceso que en el caso de un cuaderno alojado en un clúster de Amazon EMR. No es necesario configurar el cliente del usuario para poder disponer de acceso web a través de SSH, reglas de grupos de seguridad y configuraciones de proxy. Si el usuario tiene permisos suficientes, solo tiene que abrir el editor de cuadernos en la consola de Amazon EMR.

Solo un usuario puede abrir un cuaderno de EMR a la vez desde Amazon EMR. Si otro usuario intenta abrir un cuaderno de EMR que ya está abierto, se produce un error.

> **Important**
>
> Amazon EMR crea una URL prefirmada única para cada sesión del editor de cuadernos, que es válida solamente durante un breve periodo. Le recomendamos que no comparten la URL del editor de bloc de notas. De hacerlo se crea un riesgo de seguridad ya que los destinatarios de la URL adoptan sus permisos para editar el bloc de notas y ejecutar el código del bloc de notas durante la vida útil de la URL. Si otras personas necesitan acceso a un cuaderno, proporcione permisos a su usuario mediante políticas de permisos y asegúrese de que el rol de servicio de Cuadernos de EMR tenga acceso a la ubicación de Amazon S3. Para obtener más información, consulte the section called “Seguridad” y Rol de servicio para Cuadernos de Amazon EMR.

Para abrir el editor de cuadernos con un cuaderno de EMR

1. Seleccione un bloc de notas cuyo Status (Estado) sea Ready (Listo) o Pending (Pendiente) en la lista Notebooks (Blocs de notas).
2. Seleccione Abrir en o Abrir en Jupyter. JupyterLab

   Se abre una nueva pestaña del navegador en el editor JupyterLab o en el editor de Jupyter Notebook.
3. En el menú Kernel, elija Change kernel (Cambiar kernel) y, a continuación, seleccione el kernel correspondiente al lenguaje de programación que utilice.

   Ahora está preparado para escribir y ejecutar código desde el editor de blocs de notas.

Cómo guardar el contenido de un cuaderno

Cuando se trabaja en el editor de cuadernos, el contenido de las celdas del cuaderno y la salida se guardan automáticamente de forma periódica en el archivo del cuaderno en Amazon S3. Un bloc de notas que no ha sufrido cambios desde la última vez que se ha editado una celda muestra la cadena (autosaved) (guardado automático) junto al nombre del bloc de notas en el editor. Si los cambios no se han guardado todavía, aparece la cadena unsaved changes (cambios sin guardar).

Puede guardar un bloc de notas manualmente. En el menú Archivo, seleccione Guardar y punto de control o pulse CTRL+S. De este modo, se crea un archivo denominado `NotebookName.ipynb` en una carpeta de puntos de control dentro de la carpeta del cuaderno en Amazon S3. Por ejemplo, `s3://amzn-s3-demo-bucket/MyNotebookFolder/NotebookID/checkpoints/NotebookName.ipynb`. En esta ubicación, solo se guarda el archivo de punto de comprobación más reciente.

Cambio de clústeres

Puede cambiar el clúster al que está asociado un cuaderno de EMR sin cambiar el contenido del propio cuaderno. Solo se pueden cambiar de clúster los blocs de notas que tengan el estado Stopped (Detenido).

Para cambiar el clúster de un cuaderno de EMR

1. Si el bloc de notas que desea cambiar se está ejecutando, selecciónelo en la lista Notebooks (Blocs de notas) y elija Stop (Detener).

2. Cuando el estado del bloc de notas sea Stopped (Detenido), selecciónelo en la lista Notebooks (Blocs de notas) y, a continuación, elija View details (Ver detalles).
3. Elija Change cluster (Cambiar clúster).
4. Si tiene un clúster activo que ejecuta Hadoop, Spark y Livy al que desea asociar el bloc de notas, deje el valor predeterminado y seleccione un clúster en la lista. En la lista solo aparecerán los clústeres que cumplan los requisitos.

   —o—

   Elija Create a cluster (Crear clúster) y, a continuación, elija las opciones del clúster. Para obtener más información, consulte Requisitos del clúster.
5. Elija una opción para Security groups (Grupos de seguridad) y, a continuación, elija Change cluster and start notebook (Cambiar clúster e iniciar bloc de notas).

Eliminación de cuadernos y archivos de cuadernos

Cuando se elimina un cuaderno de EMR mediante la consola de Amazon EMR, se elimina el cuaderno de la lista de cuadernos disponibles. Sin embargo, los archivos de cuadernos permanecen en Amazon S3 y continúan acumulando cargos de almacenamiento.

Para eliminar un bloc de notas y sus archivos asociados

1. Abra la consola de Amazon EMR en https://console.aws.amazon.com/elasticmapreduce/.
2. Elija Notebooks (Blocs de notas), seleccione el bloc de notas en la lista y, a continuación, elija View details (Ver detalles).
3. Elija el icono de carpeta que se encuentra junto a Notebook location (Ubicación del bloc de notas) y copie la URL, que tiene el patrón `s3://MyNotebookLocationPath/NotebookID/`.
4. Elija Eliminar.

   El bloc de notas se elimina de la lista y los detalles de este ya no se pueden consultar.
5. Siga las instrucciones de ¿Cómo elimino carpetas de un bucket de S3? en la Guía del usuario de Amazon Simple Storage Service. Vaya al bucket y la carpeta del paso 3.

   —o—

   Si lo tiene AWS CLI instalado, abra una línea de comandos y escriba el comando al final de este párrafo. Sustituya la ubicación de Amazon S3 por la ubicación que ha copiado anteriormente.

Asegúrese de que AWS CLI está configurado con las claves de acceso de un usuario con permisos para eliminar la ubicación de Amazon S3. Para obtener más información, consulte Configuración de la AWS CLI en la Guía del usuario de AWS Command Line Interface .

```
aws s3 rm s3://MyNotebookLocationPath/NotebookID
```

Uso compartido de archivos de cuadernos

Cada cuaderno de EMR se guarda en Amazon S3 como un archivo denominado `NotebookName.ipynb`. Siempre que un archivo de cuaderno sea compatible con la misma versión de Cuaderno de Jupyter en la que está basado Cuadernos de EMR, puede abrir el cuaderno como un cuaderno de EMR.

La forma más sencilla de abrir un archivo de bloc de notas de otro usuario es guardar el archivo*.ipynb de otro usuario en el sistema de archivos local y, a continuación, utilizar la función de carga en Jupyter y en los editores. JupyterLab

Puede emplear este proceso para utilizar blocs de notas de EMR compartidos por otros usuarios, blocs de notas compartidos en la comunidad de Jupyter o para restaurar un bloc de notas que se ha eliminado de la consola mientras aún se conserva el archivo de bloc de notas.

Para utilizar un archivo de cuaderno diferente como base de un cuaderno de EMR

1. Antes de continuar, cierre el editor de cuadernos para los cuadernos con los que va a trabajar y, a continuación, detenga el cuaderno si se trata de un cuaderno de EMR.
2. Cree un cuaderno de EMR y asígnele un nombre. El nombre que escriba para el bloc de notas será el nombre del archivo que necesita reemplazar. El nombre de archivo nuevo debe coincidir exactamente con el nombre de este archivo.
3. Anote la ubicación en Amazon S3 que ha elegido para el cuaderno. El archivo que va a sustituir está en una carpeta con una ruta y un nombre de archivo que tienen el siguiente patrón: `s3://MyNotebookLocation/NotebookID/MyNotebookName.ipynb`.
4. Detenga el bloc de notas.
5. Sustituya el antiguo archivo de cuaderno en la ubicación de Amazon S3 por el nuevo y utilice exactamente el mismo nombre.

   El siguiente AWS CLI comando para Amazon S3 reemplaza un archivo guardado en una máquina local llamada bloc de notas EMR `SharedNotebook.ipynb` por el

nombre MyNotebook, un identificador y con el que se creó `amzn-s3-demo-bucket/MyNotebooksFolder` especificados en Amazon S3. `e-12A3BCDEFJHIJKLMNO45PQRST` Para obtener más información sobre el uso de la consola de Amazon S3 para copiar y reemplazar archivos, consulte Carga, descarga y administración de objetos en la Guía del usuario de Amazon Simple Storage Service.

```
aws s3 cp SharedNotebook.ipynb s3://amzn-s3-demo-bucket/
MyNotebooksFolder/-12A3BCDEFJHIJKLMNO45PQRST/MyNotebook.ipynb
```

Ejemplos de comandos programáticos para Cuadernos de EMR

Descripción general

Puede ejecutar cuadernos EMR ejecutándolos APIs desde un script o desde la línea de comandos. Al iniciar, detener, enumerar y describir las ejecuciones de cuadernos EMR fuera de la AWS consola, puede controlar mediante programación un cuaderno EMR. Puede pasar diferentes valores de parámetros a un cuaderno con una celda de cuaderno parametrizada. Esto elimina la necesidad de crear una copia del cuaderno para cada nuevo conjunto de valores de parámetros. Para obtener más información, consulte Acciones de la API de Amazon EMR.

Puede programar o agrupar las ejecuciones de cuadernos EMR con Amazon CloudWatch Events y. AWS Lambda Para obtener más información, consulte Uso AWS Lambda con Amazon CloudWatch Events.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Permisos de rol para la ejecución programática

Para utilizar la ejecución programática con Cuadernos de EMR, debe configurar los permisos de usuario con las siguientes políticas:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowExecutionActions",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:StartNotebookExecution",
        "elasticmapreduce:DescribeNotebookExecution",
        "elasticmapreduce:ListNotebookExecutions"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "AllowPassingServiceRole",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/EMR_Notebooks_DefaultRole"
      ]
    }
  ]
}
```

Al ejecutar mediante programación Cuadernos de EMR en un clúster de Cuadernos de EMR, debe agregar los siguientes permisos adicionales:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowRetrievingManagedEndpointCredentials",
      "Effect": "Allow",
      "Action": [
        "emr-containers:GetManagedEndpointSessionCredentials"
      ],
      "Resource": [
        "arn:aws:emr-containers:*:123456789012:/virtualclusters/virtual-cluster-
id/endpoints/managed-endpoint-id"
      ],
      "Condition": {
        "StringEquals": {
          "emr-containers:ExecutionRoleArn": [
            "arn:aws:iam::123456789012:role/emr-on-eks-execution-role"
          ]
        }
      }
    },
    {
      "Sid": "AllowDescribingManagedEndpoint",
      "Effect": "Allow",
      "Action": [
        "emr-containers:DescribeManagedEndpoint"
      ],
      "Resource": [
        "arn:aws:emr-containers:*:123456789012:/virtualclusters/virtual-cluster-
id/endpoints/managed-endpoint-id"
      ]
    }
  ]
}
```

Limitaciones de la ejecución programática

- Se admite un máximo de 100 ejecuciones simultáneas Región de AWS por cuenta.
- Una ejecución termina si dura más de 30 días.

- Las aplicaciones interactivas de Amazon EMR sin servidor no admiten la ejecución programática de cuadernos.

Ejemplos de ejecución programática de cuadernos de EMR

En las siguientes secciones se proporcionan varios ejemplos de ejecución programática de cuadernos EMR con AWS CLI el SDK de Boto3 (Python) y Ruby:

- Ejemplos de comandos de la CLI de Cuadernos en EMR Studio
- Ejemplos de Python para un cuaderno de EMR
- Ejemplos de Ruby para un cuaderno de EMR

También puede ejecutar cuadernos parametrizados como parte de los flujos de trabajo programados con una herramienta de orquestación como Apache Airflow o Amazon Managed Workflows para Apache Airflow (MWAA). Para obtener más información, consulte Orquestación de trabajos de análisis en Cuadernos de EMR mediante MWAA en el blog de macrodatos de AWS .

Ejemplos de comandos de la CLI de Cuadernos en EMR Studio

En este tema se muestran ejemplos de comandos CLI para un cuaderno de EMR. El ejemplo utiliza el cuaderno de demostración de la consola de Cuadernos de EMR. Para localizar el cuaderno, utilice la ruta del archivo relativa al directorio principal. En este ejemplo, hay dos archivos de cuadernos que puede ejecutar: `demo_pyspark.ipynb` y `my_folder/python3.ipynb`.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

La ruta relativa del archivo `demo_pyspark.ipynb` es `demo_pyspark.ipynb`, como se muestra a continuación.

La ruta relativa de `python3.ipynb` es `my_folder/python3.ipynb`, como se muestra a continuación.

Para obtener información sobre las acciones `NotebookExecution` de la API de Amazon EMR, consulte Acciones de la API de Amazon EMR.

Ejecutar un cuaderno

Puede utilizar el AWS CLI para ejecutar la `start-notebook-execution` acción en su bloc de notas, tal y como se muestra en los siguientes ejemplos.

Example — Ejecución de un cuaderno de EMR en un espacio de trabajo de EMR Studio con un clúster de Amazon EMR (que se ejecuta en Amazon) EC2

```
aws emr --region us-east-1 \
start-notebook-execution \
--editor-id e-ABCDEFG123456 \
--notebook-params '{"input_param":"my-value", "good_superhero":["superman", "batman"]}'
 \
--relative-path test.ipynb \
--notebook-execution-name my-execution \
--execution-engine '{"Id" : "j-1234ABCD123"}' \
--service-role EMR_Notebooks_DefaultRole

{
    "NotebookExecutionId": "ex-ABCDEFGHIJ1234ABCD"
}
```

Example – Ejecución de un cuaderno de EMR en un espacio de trabajo de EMR Studio con un clúster de Cuadernos de EMR

```
aws emr start-notebook-execution \
    --region us-east-1 \
    --service-role EMR_Notebooks_DefaultRole \
    --environment-variables '{"KERNEL_EXTRA_SPARK_OPTS": "--conf
 spark.executor.instances=1", "KERNEL_LAUNCH_TIMEOUT": "350"}' \
    --output-notebook-format HTML \
    --execution-engine Id=arn:aws:emr-containers:us-west-2:account-id:/
virtualclusters/ABCDEFG/
endpoints/ABCDEF,Type=EMR_ON_EKS,ExecutionRoleArn=arn:aws:iam::account-
id:role/execution-role \
```

```
    --editor-id e-ABCDEFG \
    --relative-path EMRonEKS-spark_python.ipynb
```

Example – Ejecución de un cuaderno de EMR al especificar su ubicación en Amazon S3

```
aws emr start-notebook-execution \
    --region us-east-1 \
    --notebook-execution-name my-execution-on-emr-on-eks-cluster \
    --service-role EMR_Notebooks_DefaultRole \
    --environment-variables '{"KERNEL_EXTRA_SPARK_OPTS": "--conf
 spark.executor.instances=1", "KERNEL_LAUNCH_TIMEOUT": "350"}' \
    --output-notebook-format HTML \
    --execution-engine Id=arn:aws:emr-containers:us-west-2:account-id:/
virtualclusters/ABCDEF/
endpoints/ABCDEF,Type=EMR_ON_EKS,ExecutionRoleArn=arn:aws:iam::account-
id:role/execution-role \
    --notebook-s3-location '{"Bucket": "amzn-s3-demo-bucket","Key": "s3-prefix-to-
notebook-location/EMRonEKS-spark_python.ipynb"}' \
    --output-notebook-s3-location '{"Bucket": "amzn-s3-demo-bucket","Key": "s3-prefix-
for-storing-output-notebook"}'
```

Salida de bloc de notas

Este es el resultado de un ejemplo de cuaderno. La celda 3 muestra los valores de los parámetros recién inyectados.

Describir un cuaderno

Puede utilizar la acción `describe-notebook-execution` para acceder a la información sobre la ejecución de un cuaderno específico.

```
aws emr --region us-east-1 \
describe-notebook-execution --notebook-execution-id ex-IZWZZVR9DKQ9WQ7VZWXJZR29UGHTE

{
    "NotebookExecution": {
        "NotebookExecutionId": "ex-IZWZZVR9DKQ9WQ7VZWXJZR29UGHTE",
        "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
        "ExecutionEngine": {
            "Id": "j-2QMOV6JAX1TS2",
```

```
            "Type": "EMR",
            "MasterInstanceSecurityGroupId": "sg-05ce12e58cd4f715e"
        },
        "NotebookExecutionName": "my-execution",
        "NotebookParams": "{\"input_param\":\"my-value\", \"good_superhero\":
[\"superman\", \"batman\"]}",
        "Status": "FINISHED",
        "StartTime": 1593490857.009,
        "Arn": "arn:aws:elasticmapreduce:us-east-1:123456789012:notebook-execution/ex-
IZWZZVR9DKQ9WQ7VZWXJZR29UGHTE",
        "LastStateChangeReason": "Execution is finished for cluster j-2QMOV6JAX1TS2.",
        "NotebookInstanceSecurityGroupId": "sg-0683b0a39966d4a6a",
        "Tags": []
    }
}
```

Detener un cuaderno

Si su cuaderno está ejecutando una ejecución que desea detener, puede hacerlo con el comando `stop-notebook-execution`.

```
# stop a running execution
aws emr --region us-east-1 \
stop-notebook-execution --notebook-execution-id ex-IZWZX78UVPAATC8LHJR129B1RBN4T


# describe it
aws emr --region us-east-1 \
describe-notebook-execution --notebook-execution-id ex-IZWZX78UVPAATC8LHJR129B1RBN4T

{
    "NotebookExecution": {
        "NotebookExecutionId": "ex-IZWZX78UVPAATC8LHJR129B1RBN4T",
        "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
        "ExecutionEngine": {
            "Id": "j-2QMOV6JAX1TS2",
            "Type": "EMR"
        },
        "NotebookExecutionName": "my-execution",
        "NotebookParams": "{\"input_param\":\"my-value\", \"good_superhero\":
[\"superman\", \"batman\"]}",
        "Status": "STOPPED",
        "StartTime": 1593490876.241,
```

```
        "Arn": "arn:aws:elasticmapreduce:us-east-1:123456789012:editor-execution/ex-
IZWZX78UVPAATC8LHJR129B1RBN4T",
        "LastStateChangeReason": "Execution is stopped for cluster j-2QMOV6JAX1TS2.
 Internal error",
        "Tags": []
    }
}
```

Enumerar las ejecuciones de un cuaderno por hora de inicio

Puede pasar un parámetro `--from` a `list-notebook-executions` para enumerar las ejecuciones de su cuaderno por hora de inicio.

```
# filter by start time
aws emr --region us-east-1 \
list-notebook-executions --from 1593400000.000

{
    "NotebookExecutions": [
        {
            "NotebookExecutionId": "ex-IZWZX78UVPAATC8LHJR129B1RBN4T",
            "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
            "NotebookExecutionName": "my-execution",
            "Status": "STOPPED",
            "StartTime": 1593490876.241
        },
        {
            "NotebookExecutionId": "ex-IZWZZVR9DKQ9WQ7VZWXJZR29UGHTE",
            "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
            "NotebookExecutionName": "my-execution",
            "Status": "RUNNING",
            "StartTime": 1593490857.009
        },
        {
            "NotebookExecutionId": "ex-IZWZYRS0M14L5V95WZ9OQ399SKMNW",
            "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
            "NotebookExecutionName": "my-execution",
            "Status": "STOPPED",
            "StartTime": 1593490292.995
        },
        {
            "NotebookExecutionId": "ex-IZX009ZK83IVY5E33VH8MDMELVK8K",
            "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
```

```
            "NotebookExecutionName": "my-execution",
            "Status": "FINISHED",
            "StartTime": 1593489834.765
        },
        {
            "NotebookExecutionId": "ex-IZWZXOZF88JWDF9J09GJ91R57VI0N",
            "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
            "NotebookExecutionName": "my-execution",
            "Status": "FAILED",
            "StartTime": 1593488934.688
        }
    ]
}
```

Enumerar las ejecuciones de un cuaderno por hora de inicio y estado

El comando `list-notebook-executions` también puede utilizar un parámetro `--status` para filtrar los resultados.

```
# filter by start time and status
aws emr --region us-east-1 \
list-notebook-executions --from 1593400000.000 --status FINISHED
{
    "NotebookExecutions": [
        {
            "NotebookExecutionId": "ex-IZWZZVR9DKQ9WQ7VZWXJZR29UGHTE",
            "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
            "NotebookExecutionName": "my-execution",
            "Status": "FINISHED",
            "StartTime": 1593490857.009
        },
        {
            "NotebookExecutionId": "ex-IZX009ZK83IVY5E33VH8MDMELVK8K",
            "EditorId": "e-BKTM2DIHXBEDRU44ANWRKIU8N",
            "NotebookExecutionName": "my-execution",
            "Status": "FINISHED",
            "StartTime": 1593489834.765
        }
    ]
}
```

Ejemplos de Python para un cuaderno de EMR

Este tema contiene un ejemplo de archivo de comandos. El ejemplo de código es un archivo de SDK para Python (Boto3) denominado `demo.py`. Muestra la ejecución del cuaderno. APIs

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Para obtener información sobre las acciones `NotebookExecution` de la API de Amazon EMR, consulte Acciones de la API de Amazon EMR.

```
import boto3,time

emr = boto3.client(
    'emr',
    region_name='us-west-1'
)

start_resp = emr.start_notebook_execution(
    EditorId='e-40AC8ZO6EGGCPJ4DLO48KGGGI',
    RelativePath='boto3_demo.ipynb',
    ExecutionEngine={'Id':'j-1HYZS6JQKV11Q'},
    ServiceRole='EMR_Notebooks_DefaultRole'
)

execution_id = start_resp["NotebookExecutionId"]
print(execution_id)
print("\n")

describe_response = emr.describe_notebook_execution(NotebookExecutionId=execution_id)

print(describe_response)
print("\n")
```

```
list_response = emr.list_notebook_executions()
print("Existing notebook executions:\n")
for execution in list_response['NotebookExecutions']:
    print(execution)
    print("\n")

print("Sleeping for 5 sec...")
time.sleep(5)

print("Stop execution " + execution_id)
emr.stop_notebook_execution(NotebookExecutionId=execution_id)
describe_response = emr.describe_notebook_execution(NotebookExecutionId=execution_id)
print(describe_response)
print("\n")
```

Este es el resultado de ejecutar `demo.py`.

```
ex-IZX56YJDW1D29Q1PHR32WABU2SAPK

{'NotebookExecution': {'NotebookExecutionId': 'ex-IZX56YJDW1D29Q1PHR32WABU2SAPK',
 'EditorId': 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'ExecutionEngine': {'Id':
 'j-1HYZS6JQKV11Q', 'Type': 'EMR'}, 'NotebookExecutionName': '', 'Status': 'STARTING',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 49, 19, 418000, tzinfo=tzlocal()),
 'Arn': 'arn:aws:elasticmapreduce:us-west-1:123456789012:notebook-execution/ex-
IZX56YJDW1D29Q1PHR32WABU2SAPK', 'LastStateChangeReason': 'Execution is starting
 for cluster j-1HYZS6JQKV11Q.', 'Tags': []}, 'ResponseMetadata': {'RequestId':
 '70f12c5f-1dda-45b7-adf6-964987d373b7', 'HTTPStatusCode': 200, 'HTTPHeaders': {'x-
amzn-requestid': '70f12c5f-1dda-45b7-adf6-964987d373b7', 'content-type': 'application/
x-amz-json-1.1', 'content-length': '448', 'date': 'Wed, 19 Aug 2020 00:49:22 GMT'},
 'RetryAttempts': 0}}

Existing notebook executions:

{'NotebookExecutionId': 'ex-IZX56YJDW1D29Q1PHR32WABU2SAPK', 'EditorId':
 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'NotebookExecutionName': '', 'Status': 'STARTING',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 49, 19, 418000, tzinfo=tzlocal())}


{'NotebookExecutionId': 'ex-IZX5ABS5PR1E5AHMFYEMX3JJIORRB', 'EditorId':
 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'NotebookExecutionName': '', 'Status': 'RUNNING',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 48, 36, 373000, tzinfo=tzlocal())}
```

```
{'NotebookExecutionId': 'ex-IZX5GLVXIU1HNI8BWVW057F6MF4VE', 'EditorId':
 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'NotebookExecutionName': '', 'Status': 'FINISHED',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 45, 14, 646000, tzinfo=tzlocal()),
 'EndTime': datetime.datetime(2020, 8, 19, 0, 46, 26, 543000, tzinfo=tzlocal())}


{'NotebookExecutionId': 'ex-IZX5CV8YDUO8JAIWMXN2VH32RUIT1', 'EditorId':
 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'NotebookExecutionName': '', 'Status': 'FINISHED',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 43, 5, 807000, tzinfo=tzlocal()),
 'EndTime': datetime.datetime(2020, 8, 19, 0, 44, 31, 632000, tzinfo=tzlocal())}


{'NotebookExecutionId': 'ex-IZX5AS0PPW55CEDEURZ9NSOWSUJZ6', 'EditorId':
 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'NotebookExecutionName': '', 'Status': 'FINISHED',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 42, 29, 265000, tzinfo=tzlocal()),
 'EndTime': datetime.datetime(2020, 8, 19, 0, 43, 48, 320000, tzinfo=tzlocal())}


{'NotebookExecutionId': 'ex-IZX57YF5Q53BKWLR4I5QZ14HJ7DRS', 'EditorId':
 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'NotebookExecutionName': '', 'Status': 'FINISHED',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 38, 37, 81000, tzinfo=tzlocal()),
 'EndTime': datetime.datetime(2020, 8, 19, 0, 40, 39, 646000, tzinfo=tzlocal())}

Sleeping for 5 sec...
Stop execution ex-IZX56YJDW1D29Q1PHR32WABU2SAPK
{'NotebookExecution': {'NotebookExecutionId': 'ex-IZX56YJDW1D29Q1PHR32WABU2SAPK',
 'EditorId': 'e-40AC8ZO6EGGCPJ4DLO48KGGGI', 'ExecutionEngine': {'Id':
 'j-1HYZS6JQKV11Q', 'Type': 'EMR'}, 'NotebookExecutionName': '', 'Status': 'STOPPING',
 'StartTime': datetime.datetime(2020, 8, 19, 0, 49, 19, 418000, tzinfo=tzlocal()),
 'Arn': 'arn:aws:elasticmapreduce:us-west-1:123456789012:notebook-execution/ex-
IZX56YJDW1D29Q1PHR32WABU2SAPK', 'LastStateChangeReason': 'Execution is being stopped
 for cluster j-1HYZS6JQKV11Q.', 'Tags': []}, 'ResponseMetadata': {'RequestId':
 '2a77ef73-c1c6-467c-a1d1-7204ab2f6a53', 'HTTPStatusCode': 200, 'HTTPHeaders': {'x-
amzn-requestid': '2a77ef73-c1c6-467c-a1d1-7204ab2f6a53', 'content-type': 'application/
x-amz-json-1.1', 'content-length': '453', 'date': 'Wed, 19 Aug 2020 00:49:30 GMT'},
 'RetryAttempts': 0}}
```

Ejemplos de Ruby para un cuaderno de EMR

Este tema contiene un ejemplo de Ruby que demuestra la funcionalidad de un cuaderno.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Los siguientes son ejemplos de código de Ruby que muestran el uso de la API de ejecución de cuadernos.

```
# prepare an Amazon EMR client

emr = Aws::EMR::Client.new(
  region: 'us-east-1',
  access_key_id: 'AKIA...JKPKA',
  secret_access_key: 'rLMeu...vU0OLrAC1',
)
```

Iniciar la ejecución del cuaderno y obtener el identificador de ejecución

En este ejemplo, el editor de Amazon S3 y el cuaderno de EMR son `s3://amzn-s3-demo-bucket/notebooks/e-EA8VGAA429FEQTC8HC9ZHWISK/test.ipynb`.

Para obtener información sobre las acciones `NotebookExecution` de la API de Amazon EMR, consulte Acciones de la API de Amazon EMR.

```
start_response = emr.start_notebook_execution({
    editor_id: "e-EA8VGAA429FEQTC8HC9ZHWISK",
    relative_path: "test.ipynb",

    execution_engine: {id: "j-3U82I95AMALGE"},

    service_role: "EMR_Notebooks_DefaultRole",
})


notebook_execution_id = start_resp.notebook_execution_id
```

Descripción de la ejecución del cuaderno e impresión de los detalles

```
describe_resp = emr.describe_notebook_execution({
    notebook_execution_id: notebook_execution_id
})
puts describe_resp.notebook_execution
```

El resultado de los comandos anteriores será el siguiente.

```
{
:notebook_execution_id=>"ex-IZX3VTVZWVWPP27KUB90BZ7V9IEDG",
:editor_id=>"e-EA8VGAA429FEQTC8HC9ZHWISK",
:execution_engine=>{:id=>"j-3U82I95AMALGE", :type=>"EMR", :master_instance_security_group_id=>n
:notebook_execution_name=>"",
:notebook_params=>nil,
:status=>"STARTING",
:start_time=>2020-07-23 15:07:07 -0700,
:end_time=>nil,
:arn=>"arn:aws:elasticmapreduce:us-east-1:123456789012:notebook-execution/ex-
IZX3VTVZWVWPP27KUB90BZ7V9IEDG",
:output_notebook_uri=>nil,
:last_state_change_reason=>"Execution is starting for cluster
 j-3U82I95AMALGE.", :notebook_instance_security_group_id=>nil,
:tags=>[]
}
```

Filtros para cuadernos

```
"EditorId": "e-XXXX",           [Optional]
"From" : "1593400000.000",    [Optional]
"To" :
```

Detener la ejecución del cuaderno

```
stop_resp = emr.stop_notebook_execution({
    notebook_execution_id: notebook_execution_id
})
```

Habilitar la suplantación de usuario para supervisar la actividad del usuario y del trabajo de Spark

Cuadernos de EMR le permite configurar la suplantación de usuarios en un clúster de Spark. Esta característica le ayuda a realizar un seguimiento de la actividad de los trabajos iniciados desde el propio editor de blocs de notas. Además, Cuadernos de EMR dispone de un widget de cuaderno de Jupyter para ver los detalles de los trabajos de Spark junto con la salida de las consultas en el editor de cuadernos. El widget está disponible de forma predeterminada y no requiere ninguna configuración especial. Sin embargo, para ver los servidores del historial, el cliente debe configurarse para ver las interfaces web de Amazon EMR que se alojan en el nodo principal.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Configuración de la suplantación de usuarios de Spark

De forma predeterminada, los trabajos de Spark que los usuarios envían mediante el editor de blocs de notas parecen proceder de una identidad de usuario de `livy` indefinida. Es posible configurar la suplantación de usuarios para dicho clúster para que estos trabajos se asocien a la identidad del usuario que ejecutó el código. Se crean directorios de usuario de HDFS en el nodo principal para cada identidad de usuario que ejecuta código en el cuaderno. Por ejemplo, si el usuario `NbUser1` ejecuta código desde el editor de cuadernos, puede conectarse con el nodo principal y comprobar que `hadoop fs -ls /user` muestra el directorio `/user/user_NbUser1`.

Para habilitar esta característica, establezca las propiedades de las clasificaciones de configuración `core-site` y `livy-conf`. Esta característica no está disponible de forma predeterminada si se indica a Amazon EMR que cree un clúster junto con un cuaderno. Para obtener más información sobre cómo usar clasificaciones de configuración para personalizar aplicaciones, consulte Configuración de aplicaciones en la Guía de publicación de Amazon EMR.

Utilice los siguientes valores y clasificaciones de configuración para habilitar la suplantación de usuarios en Cuadernos de EMR:

```
[
    {
        "Classification": "core-site",
        "Properties": {
          "hadoop.proxyuser.livy.groups": "*",
          "hadoop.proxyuser.livy.hosts": "*"
        }
    },
    {
        "Classification": "livy-conf",
        "Properties": {
          "livy.impersonation.enabled": "true"
        }
    }
]
```

Uso del widget de supervisión de trabajos de Spark

Cuando se ejecuta código en el editor de blocs de notas que ejecuta a su vez trabajos de Spark en el clúster de EMR, la salida incluye un widget de Jupyter Notebook para la monitorización de trabajos de Spark. El widget proporciona detalles del trabajo y enlaces útiles a la página del servidor de historial de Spark y a la página del historial de trabajos de Hadoop, junto con los enlaces correspondientes a registros de trabajos en Amazon S3 para los trabajos con errores.

Para ver las páginas del servidor de historial en el nodo principal del clúster, debe configurar un cliente SSH y un proxy de forma adecuada. Para obtener más información, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR. Para ver los registros en Amazon S3, el registro de clúster debe estar habilitado, que es la opción predeterminada para los clústeres nuevos. Para obtener más información, consulte Ver los archivos de registro archivados en Amazon S3.

A continuación, se muestra un ejemplo de supervisión de trabajos de Spark.

Control de acceso y seguridad de los cuadernos de EMR

Dispone de varias características que le pueden ayudar a personalizar el nivel de seguridad de Cuadernos de EMR. Esto ayuda a garantizar que solo los usuarios autorizados tengan acceso a

un cuaderno de EMR, puedan trabajar con los cuadernos y utilizar el editor de cuadernos para ejecutar código en el clúster. Estas características operan junto con las características de seguridad disponibles para Amazon EMR y los clústeres de Amazon EMR. Para obtener más información, consulte Seguridad en Amazon EMR.

- Puede utilizar las declaraciones AWS Identity and Access Management de política junto con las etiquetas del cuaderno para limitar el acceso. Para obtener más información, consulte Cómo funciona Amazon EMR con IAM y Ejemplo de instrucciones de políticas basadas en identidades para Cuadernos de Amazon EMR.
- Los grupos EC2 de seguridad de Amazon actúan como firewalls virtuales que controlan el tráfico de red entre la instancia principal del clúster y el editor de cuadernos. Puede utilizar los valores predeterminados o personalizar estos grupos de seguridad. Para obtener más información, consulte Especificación EC2 de grupos de seguridad para EMR Notebooks.
- Usted especifica un rol AWS de servicio que determina qué permisos tiene un bloc de notas EMR al interactuar con otros AWS servicios. Para obtener más información, consulte Rol de servicio para Cuadernos de Amazon EMR.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Instalación y uso de kernels y bibliotecas en EMR Studio

Cada cuaderno de EMR viene con un conjunto de bibliotecas y kernels preinstalados. Puede instalar bibliotecas y kernels adicionales en un clúster de EMR si el clúster tiene acceso al repositorio donde se encuentran los kernels y bibliotecas. Por ejemplo, para clústeres en subredes privadas, es posible que tenga que configurar la traducción de direcciones de red (NAT) y proporcionar una ruta para que el clúster acceda al repositorio PyPI público para instalar una biblioteca. Para obtener más información acerca de cómo configurar el acceso externo para diferentes configuraciones de red, consulte Escenarios y ejemplos en la Guía del usuario de Amazon VPC.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Las aplicaciones EMR Serverless vienen con las siguientes bibliotecas preinstaladas para Python y: PySpark

- Bibliotecas Python: ggplot, matplotlib, numpy, pandas, plotly, bokeh, scikit-learn, scipy, scipy
- PySpark bibliotecas —ggplot,,matplotlib,numpy,,pandas,plotly, bokeh scikit-learn scipy scipy

Instalación de kernels y bibliotecas de Python en un nodo principal del clúster

Con la versión 5.30.0 de Amazon EMR y posteriores, excepto 6.0.0, puede instalar bibliotecas y kernels de Python adicionales en el nodo principal del clúster. Después de la instalación, estos kernels y bibliotecas están disponibles para cualquier usuario que ejecute un cuaderno de EMR asociado al clúster. Las bibliotecas de Python instaladas de esta manera solo están disponibles para procesos que se ejecutan en el nodo principal. Las bibliotecas no están instaladas en nodos principales o de tareas y no están disponibles para los ejecutores que se ejecutan en esos nodos.

> **Note**
>
> Para las versiones 5.30.1, 5.31.0 y 6.1.0 de Amazon EMR, debe tomar medidas adicionales para instalar los kernels y las bibliotecas en el nodo principal de un clúster.
> Para habilitar la característica, haga lo siguiente:
>
> 1. Asegúrese de que la política de permisos adjunta al rol de servicio de Cuadernos de EMR permita la siguiente acción:
>
>  ```
>  elasticmapreduce:ListSteps
>  ```

Para obtener más información, consulte Rol de servicio para Cuadernos de EMR.

2. Utilice el AWS CLI para ejecutar un paso en el clúster que configura los EMR Notebooks, como se muestra en el siguiente ejemplo. Debe usar el nombre del paso `EMRNotebooksSetup`. *us-east-1*Sustitúyalo por la región en la que reside el clúster. Para obtener más información, consulte Adición de pasos a un clúster mediante la AWS CLI.

```
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"]
```

Puede instalar kernels y bibliotecas con `pip` o `conda` en el directorio `/emr/notebook-env/bin` del nodo principal.

Example – Instalación de bibliotecas de Python

Desde el kernel de Python3, ejecute la magia de `%pip` como un comando desde el interior de una celda de un cuaderno para instalar las bibliotecas de Python.

```
%pip install pmdarima
```

Es posible que necesite reiniciar el kernel para usar los paquetes actualizados. También puede usar la magia de Spark de `%%sh` para invocar `pip`.

```
%%sh
/emr/notebook-env/bin/pip install -U matplotlib
/emr/notebook-env/bin/pip install -U pmdarima
```

Si utiliza un PySpark núcleo, puede instalar bibliotecas en el clúster mediante `pip` comandos o utilizar bibliotecas integradas en un bloc de notas. PySpark

Para ejecutar comandos `pip` en el clúster desde la terminal, conéctese primero al nodo principal mediante SSH, tal como se muestra en los siguientes comandos.

```
sudo pip3 install -U matplotlib
```

```
sudo pip3 install -U pmdarima
```

Como alternativa, puede utilizar bibliotecas adaptadas a un cuaderno. En el caso de las bibliotecas adaptadas a un cuaderno, la instalación de la biblioteca se limita al ámbito de la sesión y se hace en todos los ejecutores de Spark. Para obtener más información, consulte Uso de bibliotecas adaptadas a los cuadernos.

Si desea empaquetar varias bibliotecas de Python en un PySpark núcleo, también puede crear un entorno virtual de Python aislado. Para ver ejemplos, consulte Using Virtualenv.

Para crear un entorno virtual de Python en una sesión, utilice la propiedad `spark.yarn.dist.archives` de Spark del comando mágico `%%configure` en la primera celda de un cuaderno, tal como se muestra en el siguiente ejemplo.

```
%%configure -f
{
   "conf": {
   "spark.yarn.appMasterEnv.PYSPARK_PYTHON":"./environment/bin/python",
   "spark.yarn.appMasterEnv.PYSPARK_DRIVER_PYTHON":"./environment/bin/python",
   "spark.yarn.dist.archives":"s3://amzn-s3-demo-bucket/prefix/
my_pyspark_venv.tar.gz#environment",
   "spark.submit.deployMode":"cluster"
   }
}
```

De manera similar, puede crear un entorno ejecutor de Spark.

```
%%configure -f
{
   "conf": {
   "spark.yarn.appMasterEnv.PYSPARK_PYTHON":"./environment/bin/python",
   "spark.yarn.appMasterEnv.PYSPARK_DRIVER_PYTHON":"./environment/bin/python",
   "spark.executorEnv.PYSPARK_PYTHON":"./environment/bin/python",
   "spark.yarn.dist.archives":"s3://amzn-s3-demo-bucket/prefix/
my_pyspark_venv.tar.gz#environment",
   "spark.submit.deployMode":"cluster"
   }
}
```

También puede utilizar `conda` para instalar bibliotecas de Python. No necesita tener acceso a sudo para poder utilizar `conda`. Debe conectarse al nodo principal mediante SSH y luego ejecutar `conda`

desde el terminal. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.

Example – Instalación de kernels

El siguiente ejemplo muestra cómo instalar el kernel de Kotlin mediante un comando de terminal mientras está conectado al nodo principal de un clúster:

```
sudo /emr/notebook-env/bin/conda install kotlin-jupyter-kernel -c jetbrains
```

> **Note**
>
> Estas instrucciones no instalan las dependencias del kernel. Si su kernel tiene dependencias de terceros, es posible que deba seguir pasos de configuración adicionales antes de poder utilizar el kernel con su cuaderno.

Consideraciones y limitaciones de las bibliotecas adaptadas al cuaderno

Tenga en cuenta lo siguiente al utilizar bibliotecas adaptadas al cuaderno:

- Las bibliotecas adaptadas al cuaderno están disponibles para los clústeres que cree con las versiones 5.26.0 y posteriores de Amazon EMR.
- Las bibliotecas integradas en ordenadores portátiles están pensadas para usarse únicamente con el núcleo. PySpark
- Cualquier usuario puede instalar bibliotecas adaptadas al bloc de notas adicionales desde una celda de bloc de notas. Estas bibliotecas solo están disponibles para ese usuario de bloc de notas durante una sola sesión de bloc de notas. Si otros usuarios necesitan las mismas bibliotecas o el mismo usuario necesita las mismas bibliotecas en una sesión diferente, la biblioteca debe volver a instalarse.
- Solo puede desinstalar las bibliotecas que se hayan instalado con la API de `install_pypi_package`. No puede desinstalar ninguna biblioteca que esté preinstalada en el clúster.
- Si se instalan las mismas bibliotecas con diferentes versiones en el clúster y como bibliotecas adaptadas al bloc de notas, la versión de la biblioteca con ámbito de bloc de notas anula la versión de la biblioteca del clúster.

Uso de bibliotecas adaptadas al cuaderno

Para instalar bibliotecas, el clúster de Amazon EMR debe tener acceso al repositorio PyPI en el que se encuentran las bibliotecas.

Los siguientes ejemplos muestran comandos sencillos para enumerar, instalar y desinstalar bibliotecas desde el interior de una celda de un portátil mediante el núcleo y. PySpark APIs Para ver ejemplos adicionales, consulte la publicación Instalar bibliotecas de Python en un clúster en ejecución con EMR Notebooks en AWS el blog Big Data.

Example – Enumeración de bibliotecas actuales

El siguiente comando muestra los paquetes de Python disponibles para la sesión de bloc de notas de Spark actual. Esto muestra las bibliotecas instaladas en el clúster y las bibliotecas adaptadas al bloc de notas.

```
sc.list_packages()
```

Example – Instalación de la biblioteca de Celery

El siguiente comando instala la biblioteca de Celery como una biblioteca adaptada al bloc de notas.

```
sc.install_pypi_package("celery")
```

Después de instalar la biblioteca, el siguiente comando confirma que la biblioteca está disponible en el controlador Spark y los ejecutores.

```
import celery
sc.range(1,10000,1,100).map(lambda x: celery.__version__).collect()
```

Example – Instalación de la biblioteca de Arrow, con la versión y el repositorio especificados

El siguiente comando instala la biblioteca Arrow como una biblioteca adaptada al bloc de notas, con una especificación de la versión de la biblioteca y la URL del repositorio.

```
sc.install_pypi_package("arrow==0.14.0", "https://pypi.org/simple")
```

Example – Desinstalación de una biblioteca

El siguiente comando desinstala la biblioteca Arrow y la elimina de la sesión actual como biblioteca adaptada al bloc de notas.

```
sc.uninstall_package("arrow")
```

Asociación de repositorios basados en Git con Cuadernos de EMR

Puede asociar repositorios basados en Git con sus cuadernos de Amazon EMR para guardar los cuadernos en un entorno con control de versiones. Puede asociar hasta tres repositorios a un bloc de notas. Se admiten los siguientes servicios basados en Git:

- AWS CodeCommit
- GitHub
- Bitbucket
- GitLab

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

La asociación de repositorios basados en Git con su bloc de notas tiene las siguientes ventajas.

- Control de versiones: puede registrar los cambios de código en un sistema de control de versiones para que pueda revisar el historial de sus cambios y revertir selectivamente algunos de ellos.
- Colaboración: los compañeros que trabajan en diferentes cuadernos pueden compartir código a través de repositorios remotos basados en Git. Puede clonar o combinar código de repositorios remotos e insertar los cambios en esos repositorios remotos.
- Reutilización de código: muchos cuadernos de Jupyter que muestran técnicas de análisis de datos o aprendizaje automático están disponibles en repositorios alojados públicamente, como. GitHub Puede asociar sus blocs de notas a un repositorio para reutilizar los blocs de notas de Jupyter incluidos en un repositorio.

Para usar repositorios basados en Git con Cuadernos de EMR, agregue los repositorios como recursos en la consola de Amazon EMR, asocie credenciales a los repositorios que requieran autenticación y vincule los repositorios con sus cuadernos. Puede ver una lista de repositorios que se almacenan en su cuenta y obtener más información sobre cada repositorio en la consola de Amazon EMR. Puede asociar un repositorio basado en Git existente con un bloc de notas al crearlo.

Temas

- Requisitos previos y consideraciones al integrar un cuaderno EMR con un repositorio
- Agregar un repositorio basado en Git a Amazon EMR
- Actualizar o eliminar un repositorio basado en Git de un espacio de trabajo de EMR Studio
- Vincular o desvincular un repositorio basado en Git en EMR Studio
- Crear un nuevo cuaderno con un repositorio de Git asociado en EMR Studio
- Uso de repositorios de Git en un Cuaderno de EMR Studio

Requisitos previos y consideraciones al integrar un cuaderno EMR con un repositorio

Tenga en cuenta las siguientes buenas prácticas con respecto a las confirmaciones, permisos y hosting al planear la integración de un repositorio basado en Git con Cuadernos de EMR.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

AWS CodeCommit

Si utilizas un CodeCommit repositorio, debes usar las credenciales de Git y HTTPS con CodeCommit. No se admiten las claves SSH ni HTTPS con el asistente de AWS CLI credenciales. CodeCommit no admite los tokens de acceso personal (). PATs Para obtener más información, consulte Uso de IAM con CodeCommit: credenciales de Git, claves SSH y claves de AWS acceso en

la Guía del usuario de IAM y Configuración para usuarios de HTTPS que usan credenciales de Git en la Guía del AWS CodeCommit usuario.

Consideraciones de acceso y permisos

Antes de asociar un repositorio a su cuaderno, debe asegurarse de que el clúster, el rol de IAM de Cuadernos de EMR y los grupos de seguridad tengan la configuración y los permisos correctos. También puede configurar los repositorios basados en Git que aloje en una red privada siguiendo las instrucciones que se indican en Configurar un repositorio Git alojado de forma privada para Cuadernos de EMR.

- Acceso a Internet del clúster: la interfaz de red que se inicia solo tiene una dirección IP privada. Esto significa que el clúster al que se conecta el bloc de notas debe estar en una subred privada con una gateway de traducción de direcciones de red (NAT) o debe poder obtener acceso a Internet a través de una gateway privada virtual. Para obtener más información, consulte Opciones de Amazon VPC.

  Los grupos de seguridad del bloc de notas deben incluir una regla de salida que permita que el bloc de notas envíe tráfico a Internet desde el clúster. Le recomendamos que cree sus propios grupos de seguridad. Para obtener más información, consulte Especificación de grupos EC2 de seguridad para los EMR Notebooks.

  > **Important**
  >
  > Si la interfaz de red se inicia en una subred pública, no podrá comunicarse con Internet a través de una puerta de enlace de Internet (IGW).

- Permisos para AWS Secrets Manager: si utilizas Secrets Manager para almacenar los secretos que utilizas para acceder a un repositorio, the section called “Rol de Cuadernos de Amazon EMR” debe tener una política de permisos adjunta que permita la `secretsmanager:GetSecretValue` acción.

Configurar un repositorio Git alojado de forma privada para Cuadernos de EMR

Utilice las siguientes instrucciones para configurar los repositorios alojados de forma privada para Cuadernos de EMR. Debe proporcionar un archivo de configuración con información sobre sus servidores de DNS y Git. Amazon EMR utiliza esta información para configurar cuadernos de EMR que puedan dirigir el tráfico a sus repositorios alojados de forma privada.

Requisitos previos

Para obtener acceso a un repositorio de Git alojado de forma privada para Cuadernos de EMR, debe tener lo siguiente:

- Una Amazon S3 Control ubicación en la que se guardarán los archivos de su bloc de notas EMR.

Para configurar uno o más repositorios de Git alojados de forma privada para Cuadernos de EMR

1. Cree un archivo de configuración con la plantilla proporcionada. Incluya los siguientes valores para cada servidor de Git que desee especificar en la configuración:

   - **`DnsServerIpV4`**- La IPv4 dirección de su servidor DNS. Si proporciona valores para `DnsServerIpV4` y `GitServerIpV4List`, el valor de `DnsServerIpV4` tiene prioridad y se utilizará para resolver el `GitServerDnsName`.

     > **Note**
     > Para usar repositorios de Git alojados de forma privada, su servidor de DNS debe permitir el acceso entrante desde Cuadernos de EMR. Le recomendamos que proteja su servidor de DNS contra otros accesos no autorizados.

   - **`GitServerDnsName`**: el nombre de DNS del servidor de Git. Por ejemplo, `"git.example.com"`.
   - **`GitServerIpV4List`**- Una lista de IPv4 direcciones que pertenecen a tus servidores Git.

```
[
    {
        "Type": "PrivatelyHostedGitConfig",
        "Value": [
            {
                "DnsServerIpV4": "<10.24.34.xxx>",
                "GitServerDnsName": "<enterprise.git.com>",
                "GitServerIpV4List": [
                    "<xxx.xxx.xxx.xxx>",
                    "<xxx.xxx.xxx.xxx>"
                ]
            },
            {
```

```
                "DnsServerIpV4": "<10.24.34.xxx>",
                "GitServerDnsName": "<git.example.com>",
                "GitServerIpV4List": [
                    "<xxx.xxx.xxx.xxx>",
                    "<xxx.xxx.xxx.xxx>"
                ]
            }
        ]
    }
]
```

2. Guarde su archivo de configuración como `configuration.json`.
3. Cargue el archivo de configuración en la ubicación de almacenamiento designada de Amazon S3, en una carpeta llamada `life-cycle-configuration`. Por ejemplo, si su ubicación de S3 predeterminada es `s3://amzn-s3-demo-bucket/notebooks`, el archivo de configuración debe estar ubicado en `s3://amzn-s3-demo-bucket/notebooks/`**`life-cycle-configuration/configuration.json`**.

> **Important**
>
> Le recomendamos que restrinja el acceso a su carpeta `life-cycle-configuration` únicamente a los administradores de Cuadernos de EMR y al rol de servicio de Cuadernos de EMR. También debe protegerse contra el acceso no autorizado de `configuration.json`. Para obtener instrucciones, consulte Controlar el acceso a un bucket con políticas de usuario o Prácticas recomendadas de seguridad para Amazon S3.

Para ver las instrucciones de carga, consulte Creación de una carpeta y Carga de objetos en la Guía del usuario de Amazon Simple Storage Service.

Agregar un repositorio basado en Git a Amazon EMR

Consulte las siguientes secciones para obtener información sobre cómo agregar un repositorio basado en Git a un cuaderno de EMR de la consola antigua o a un Espacio de trabajo de EMR Studio de la consola.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Console

Como Cuadernos de EMR es Espacios de trabajo de EMR Studio en la nueva consola, puede seguir las instrucciones de Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio para asociar hasta tres repositorios de Git a su espacio de trabajo.

También puede utilizar la extensión JupyterLab Git. Seleccione el icono de Git en la barra lateral izquierda de su cuaderno de JupyterLab para acceder a la extensión. Para obtener información sobre la extensión, consulta el repositorio GitHub jupyterlab-git.

Para asociar un repositorio de Git a un espacio de trabajo, el administrador de Studio debe tomar medidas para configurar Studio y permitir la vinculación de repositorios de Git. Para obtener más información, consulte Establecer el acceso y los permisos para los repositorios basados en Git.

Actualizar o eliminar un repositorio basado en Git de un espacio de trabajo de EMR Studio

Consulte las siguientes secciones para obtener información sobre cómo eliminar un repositorio basado en Git de un cuaderno de EMR de la consola antigua o de un Espacio de trabajo de EMR Studio de la consola.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la nueva consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte

Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Console

Como Cuadernos de EMR es Espacios de trabajo de EMR Studio en la nueva consola, puede consultar Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio para obtener más información sobre cómo trabajar con los repositorios de Git en su espacio de trabajo. Sin embargo, en este momento, no puede eliminar los repositorios de Git de los espacios de trabajo.

Vincular o desvincular un repositorio basado en Git en EMR Studio

Siga estos pasos para vincular o desvincular un repositorio basado en Git a un cuaderno de EMR de la consola anterior o a un Espacio de trabajo de EMR Studio de la consola.

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Console

Como Cuadernos de EMR es Espacios de trabajo de EMR Studio en la nueva consola, puede consultar Vincular repositorios basados en Git a un espacio de trabajo de EMR Studio para obtener más información sobre cómo trabajar con los repositorios de Git en su espacio de trabajo. Sin embargo, en este momento, no puede eliminar los repositorios de Git de los espacios de trabajo.

Descripción del estado del repositorio

Un repositorio Git puede tener cualquiera de los siguientes estados en la lista de repositorios. Para obtener más información sobre cómo vincular EMR Notebooks con los repositorios Git, consulte Vincular o desvincular un repositorio basado en Git en EMR Studio.

Estado	Significado
Linking (En proceso de vinculación)	El repositorio de Git se está vinculando al bloc de notas. Mientras el repositorio tiene el estado Linking (En proceso de vinculación), no se puede detener el bloc de notas.
Linked (Vinculado)	El repositorio de Git está vinculado al bloc de notas. Mientras el repositorio tenga el estado Linked (Vinculado), estará conectado al repositorio remoto.
Link Failed (Error al vincular)	El repositorio de Git no pudo vincularse al bloc de notas. Puede intentar vincularlo de nuevo.
Unlinking (En proceso de desvinculación)	El repositorio de Git se está desvinculando del bloc de notas. Mientras el repositorio tenga el estado Unlinking (En proceso de desvincul ación), no podrá detener el bloc de notas. Cuando se desvincula un repositorio de Git de un bloc de notas solo se desconecta del repositorio remoto: no se elimina ningún código del bloc de notas.
Unlink Failed (Error al desvincular)	El repositorio de Git no se pudo desvincular del bloc de notas. Puede intentar desvincularlo de nuevo.

Crear un nuevo cuaderno con un repositorio de Git asociado en EMR Studio

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Para crear un cuaderno y asociarlo con repositorios de Git en la consola de Amazon EMR antigua

1. Siga las instrucciones de Crear un cuaderno en EMR Studio.
2. En Security group (Grupo de seguridad), elija Use your own security group (Usar su propio grupo de seguridad).

   > **Note**
   >
   > Los grupos de seguridad del bloc de notas deben incluir una regla de salida para permitir que el bloc de notas envíe tráfico a Internet a través del clúster. Le recomendamos que cree sus propios grupos de seguridad. Para obtener más información, consulte Especificación de grupos EC2 de seguridad para los EMR Notebooks.

3. En Git repositories (Repositorios de Git), elija el repositorio que desee asociar al bloc de notas.
   1. Elija un repositorio almacenado como recurso en su cuenta y, a continuación, elija Save (Guardar).
   2. Para añadir un nuevo repositorio como recurso a su cuenta, elija add a new repository (añadir un nuevo repositorio). Complete el flujo de trabajo Add repository (Añadir repositorio) en una ventana nueva.

Uso de repositorios de Git en un Cuaderno de EMR Studio

> **Note**
>
> Cuadernos de EMR está disponible como Espacios de trabajo de EMR Studio en la consola. El botón Crear espacio de trabajo de la consola le permite crear nuevos cuadernos. Para crear espacios de trabajo o acceder a ellos, los usuarios de Cuadernos de EMR necesitan permisos de rol de IAM adicionales. Para obtener más información, consulte Cuadernos de Amazon EMR es Espacios de trabajo de Amazon EMR Studio en la consola y Consola Amazon EMR.

Al abrir un bloc de notas, puede elegir entre Abrir en Jupyter JupyterLab o Abrir en Jupyter.

Si elige abrir el bloc de notas en Jupyter, se muestra una lista expandible de archivos y carpetas dentro del bloc de notas. Puede ejecutar manualmente comandos de Git como los siguientes en una celda del bloc de notas.

```
!git pull origin primary
```

Para abrir cualquiera de los repositorios adicionales, desplácese hasta otras carpetas.

Si elige abrir el cuaderno con una JupyterLab interfaz, puede usar la extensión JupyterLab Git preinstalada. Para obtener más información acerca de la extensión, consulte jupyterlab-git.

Planificación, configuración y lanzamiento de clústeres de Amazon EMR

En esta sección se explican las opciones de configuración y las instrucciones de planificación, configuración y lanzamiento de clústeres con Amazon EMR. Antes de lanzar un clúster, elija opciones del sistema en función de los datos que está procesando y de los requisitos de costo, velocidad, capacidad, disponibilidad, seguridad y capacidad de administración. Las opciones incluyen:

- La región donde ejecutar un clúster, dónde y cómo almacenar datos y cómo extraer resultados. Consulte Configuración de la ubicación del clúster de Amazon EMR y el almacenamiento de datos.
- Si está ejecutando clústeres de Amazon EMR en Outposts o zonas locales. Consulte EMR se agrupa en AWS Outposts o Clústeres de EMR en Zonas Locales AWS.
- Si un clúster es de ejecución prolongada o transitorio y el software que ejecuta. Consulte Configuración de un clúster de Amazon EMR para que continúe o finalice después de la ejecución de pasos y Configure las aplicaciones al lanzar su clúster de Amazon EMR.
- Si un clúster tiene un solo nodo principal o tres nodos principales. Consulte Planificación y configuración de nodos principales en su clúster de Amazon EMR.
- Las opciones de hardware y red que optimizan costos, el desempeño y la disponibilidad para la aplicación. Consulte Configuración del hardware y las redes de los clústeres de Amazon EMR.
- Cómo configurar los clústeres para que pueda administrarlos con más facilidad y monitorizar la actividad, el desempeño y la salud. Consulte Configuración del registro y la depuración de un clúster de Amazon EMR y Etiquetado y clasificación de los recursos del clúster de Amazon EMR.
- Cómo autenticar y autorizar el acceso a los recursos de clústeres y cómo cifrar los datos. Consulte Seguridad en Amazon EMR.
- Cómo integrarlo con otro software y servicios. Consulte Integración de controladores y aplicaciones de terceros en Amazon EMR.

Lanzamiento de un clúster de Amazon EMR

Siga estos pasos para lanzar un clúster de Amazon EMR en tan solo unos minutos.

Para lanzar un clúster rápidamente utilizando la consola

1. Inicie sesión en la AWS Management Console consola Amazon EMR y ábrala en https://console.aws.amazon.com/emr/ los clústeres.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En la página Crear clúster, ingrese o seleccione valores para los campos proporcionados. En el panel de resumen persistente se muestra una vista en tiempo real de las opciones de clúster seleccionadas actualmente. Seleccione un encabezado en el panel de resumen para ir a la sección correspondiente y realizar los ajustes. El nombre del clúster no puede contener los caracteres <, >, $, | ni ` (acento grave). Debe completar todas las configuraciones requeridas antes de seleccionar Crear clúster.
4. Seleccione Crear clúster para aceptar la configuración tal como se muestra.
5. Se abrirá la página de detalles del clúster. Consulte el estado del clúster junto a su nombre. El estado debería cambiar de Iniciando a En ejecución y a Esperando durante el proceso de creación del clúster. Es posible que tenga que elegir el icono de actualización situado en la parte superior derecha o actualizar el navegador para recibir las actualizaciones.

   Cuando el estado cambia a Esperando, el clúster está activo, ejecutándose y listo para aceptar pasos y conexiones SSH.

Configuración de la ubicación del clúster de Amazon EMR y el almacenamiento de datos

Esta sección describe cómo configurar la región de un clúster, los diferentes sistemas de archivos disponibles cuando utilice Amazon EMR y cómo utilizarlos. También aborda la preparación o carga de datos en Amazon EMR si es necesario, así como la manera de preparar una ubicación de salida para archivos de registro y cualquier archivo de datos de salida que configure.

Temas

- Elija una AWS región para su clúster de Amazon EMR
- Trabajo con almacenamiento y sistemas de archivos con Amazon EMR
- Preparación de los datos de entrada para su procesamiento con Amazon EMR
- Cómo configurar una ubicación para la salida del clúster de Amazon EMR

Elija una AWS región para su clúster de Amazon EMR

Amazon Web Services se ejecuta en los servidores de los centros de datos de todo el mundo. Los centros de datos están organizados en regiones geográficas. Al lanzar un clúster de Amazon EMR, debe especificar una región. Puede elegir la región para reducir la latencia, minimizar los costos o satisfacer los requisitos reglamentarios. Para ver la lista de regiones y puntos de conexión que admite Amazon EMR, consulte Regiones y puntos de conexión en la Referencia general de Amazon Web Services.

Para obtener el máximo rendimiento, debe lanzar el clúster en la misma región donde se encuentren sus datos. Por ejemplo, si el bucket de Amazon S3 que almacena sus datos de entrada se encuentra en la región Oeste de EE. UU. (Oregón), debe lanzar el clúster en esa misma región para evitar las tarifas de transferencia de datos entre regiones. Si utiliza un bucket de Amazon S3 para recibir la salida del clúster, también debería crearlo en la región Oeste de EE. UU. (Oregón).

Si planea asociar un par de EC2 claves de Amazon al clúster (necesario para usar SSH para iniciar sesión en el nodo principal), el par de claves debe crearse en la misma región que el clúster. Del mismo modo, los grupos de seguridad que crea Amazon EMR para administrar el clúster se crean en la misma región que este.

Si te registraste Cuenta de AWS en una fecha posterior al 17 de mayo de 2017, la región predeterminada para acceder a un recurso desde allí AWS Management Console es EE.UU. Este (Ohio) (us-east-2); para las cuentas más antiguas, la región predeterminada es EE.UU. Oeste (Oregón) (us-west-2) o EE.UU. Este (Norte de Virginia) (us-east-1). Para obtener más información, consulte Regiones y puntos de conexión.

Algunas AWS funciones solo están disponibles en regiones limitadas. Por ejemplo, las instancias de informática del clúster solo están disponibles en la región Este de EE. UU. (Norte de Virginia) y la región Asia-Pacífico (Sídney) admite solo Hadoop 1.0.3 y versiones posteriores. Cuando elija una región, compruebe que admite las características que desea utilizar.

Para obtener el mejor rendimiento, utilice la misma región para todos los AWS recursos que vaya a utilizar con el clúster. La siguiente tabla asigna los nombres de región entre servicios. Para obtener una lista de las regiones de Amazon EMR, consulte Regiones de AWS y puntos de conexión en la Referencia general de Amazon Web Services.

Elegir una región mediante la consola

La región predeterminada aparece del lado izquierdo de la información de la cuenta en la barra de navegación. Para cambiar de región tanto en la consola nueva como en la antigua, seleccione el menú desplegable Región y seleccione una nueva opción.

Especifique una región con AWS CLI

Especifique una región por defecto AWS CLI mediante el aws configure comando o la variable de `AWS_DEFAULT_REGION` entorno. Para obtener más información, consulte Configuración de la AWS región en la Guía del AWS Command Line Interface usuario.

Elegir una región mediante un SDK o la API

Para elegir una región con un SDK, configure su aplicación para que utilice el punto de conexión de esa región. Si está creando una aplicación cliente mediante un SDK de AWS , puede cambiar el punto de conexión del cliente llamando `setEndpoint`, tal y como se muestra en el ejemplo siguiente:

```
client.setEndpoint("elasticmapreduce.us-west-2.amazonaws.com");
```

Una vez que la aplicación haya especificado una región mediante la configuración del punto final, puede configurar la zona de disponibilidad para EC2 las instancias del clúster. Las zonas de disponibilidad son regiones geográficas diferentes que se han diseñado para estar aisladas de errores que se produzcan en otras zonas de disponibilidad y que proporcionan conectividad de red de baja latencia a otras zonas de disponibilidad de la misma región. Una región consta de una o varias zonas de disponibilidad. Para optimizar el rendimiento y reducir la latencia, todos los recursos deben estar situados en la misma zona de disponibilidad que el clúster que los utiliza.

Trabajo con almacenamiento y sistemas de archivos con Amazon EMR

Amazon EMR y Hadoop proporcionan una variedad de sistemas de archivos que puede utilizar al momento de procesar pasos de clústeres. Puede especificar qué sistema de archivos utilizar mediante el prefijo del URI utilizado para acceder a los datos. Por ejemplo, `s3://amzn-s3-demo-bucket1/path` hace referencia a un bucket de Amazon S3 mediante S3A (desde la versión EMR-7.10.0). En la siguiente tabla se muestran los sistemas de archivos disponibles, con recomendaciones sobre la hora que es mejor utilizar cada una de ellas.

Amazon EMR y Hadoop suelen utilizar dos o más de los siguientes sistemas de archivos al procesar un clúster. HDFS y S3A son los dos sistemas de archivos principales que se utilizan con Amazon EMR.

> **Important**
>
> A partir de la versión 5.22.0 de Amazon EMR, Amazon EMR AWS utiliza Signature Version 4 exclusivamente para autenticar las solicitudes a Amazon S3. Las versiones anteriores de Amazon EMR utilizan AWS Signature Version 2 en algunos casos, a menos que las notas de la versión indiquen que se utiliza exclusivamente Signature Version 4. Para obtener más información, consulte Autenticación de solicitudes (versión de AWS firma 4) y Autenticación de solicitudes (versión de AWS firma 2) en la Guía para desarrolladores de Amazon Simple Storage Service.

Sistema de archivos	Prefijo	Descripción
HDFS	`hdfs://` (o sin prefijo)	HDFS es un sistema de archivos distribuido, escalable y portátil para Hadoop. Una ventaja de HDFS es el reconocimiento de datos entre los nodos de clúster de Hadoop que administran los clústeres y los nodos de clúster de Hadoop que administran los pasos individuales. Para obtener más información, consulte la documentación de Hadoop. Los nodos principales y los nodos secundarios utilizan HDFS. Una ventaja es que es rápido; una desventaj a es que se trata de almacenamiento efímero que se reclama cuando el clúster finaliza. Es mejor utilizarlo para almacenar en caché los resultados producidos por pasos de flujos de trabajo intermedios.
S3A	`s3://`, `s3a://`, `s3n://`	El sistema de archivos Hadoop S3A es un conector S3 de código abierto que permite a Apache Hadoop y su ecosistema interactuar directamente con el almacenamiento de Amazon S3. Permite a los usuarios leer y escribir datos en depósitos de S3

Sistema de archivos	Prefijo	Descripción
		mediante operaciones de archivos compatibles con Hadoop, lo que proporciona una integración perfecta entre las aplicaciones de Hadoop y el almacenamiento en la nube. **Note** Antes de EMR-7.10.0, Amazon EMR utilizaba EMRFS para los esquemas s3://y s3n://.
sistema de archivos local		El sistema de archivos local se refiere a un disco conectado a nivel local. Cuando se crea un clúster de Hadoop, cada nodo se crea a partir de una EC2 instancia que incluye un bloque preconfigurado de almacenamiento en disco preconectado denominad o almacén de instancias. Los datos de los volúmenes del almacén de instancias solo se conservan durante la vida útil de la instancia. EC2 Los volúmenes de almacén de instancias son ideales para el almacenam iento temporal de datos que cambian constantemente, como los búferes, las cachés, los datos de pruebas y otro contenido temporal. Para obtener más informaci ón, consulta Amazon EC2 instance Storage. HDFS utiliza el sistema de archivos local, pero Python también se ejecuta desde el sistema de archivos local y puede optar por almacenar archivos de aplicaciones adicionales en volúmenes de almacenes de instancias.

Sistema de archivos	Prefijo	Descripción
Sistema de archivos de bloques de Amazon S3 (heredado)	`s3bfs://`	El sistema de archivos de bloque de Amazon S3 es un sistema de almacenamiento de archivos heredado. Recomendamos encarecidamente evitar el uso de este sistema. **Important** Le recomendamos que no use este sistema de archivos, ya que puede activar una condición de carrera que podrían provocar un error del clúster. Sin embargo, es posible que aplicacio nes heredadas lo requieran.

Acceso a sistemas de archivo

Puede especificar qué sistema de archivos utilizar mediante el prefijo del identificador de recursos uniforme (URI) utilizado para acceder a los datos. Los siguientes procedimientos ilustran cómo hacer referencia a diferentes tipos de sistemas de archivos.

Para acceder a una HDFS local

- Especifique el prefijo `hdfs:///` en el URI. Amazon EMR resuelve rutas que no especifican un prefijo en el URI al HDFS local. Por ejemplo, las dos situaciones siguientes se URIs resolverían en la misma ubicación en HDFS.

```
hdfs:///path-to-data

/path-to-data
```

Para acceder a una HDFS remota

- Incluya la dirección IP del nodo principal en el URI, tal y como se muestra en los siguientes ejemplos.

```
hdfs://master-ip-address/path-to-data

master-ip-address/path-to-data
```

Para acceder a Amazon S3

- Utilice el prefijo `s3://`.

```
s3://bucket-name/path-to-file-in-bucket
```

Para acceder al sistema de archivos de bloques de Amazon S3

- Se utiliza únicamente por las aplicaciones heredadas que requieren el sistema de archivos de bloques de Amazon S3. Para acceder a datos o almacenar datos con este sistema de archivos, utilice el prefijo `s3bfs://` en el URI.

  El sistema de archivos en bloque Amazon S3 es un sistema de archivos heredado que se utilizaba para admitir cargas a Amazon S3 de más de 5 GB. Con la funcionalidad de carga multiparte que Amazon EMR proporciona a través AWS del SDK de Java, puede cargar archivos de gran tamaño al sistema de archivos nativo de Amazon S3, y el sistema de archivos de bloques de Amazon S3 está obsoleto. Para obtener más información sobre la carga multiparte para EMR, consulte Configurar la carga multiparte para Amazon S3. Para obtener más información sobre los límites de tamaño de objetos y piezas de S3, consulte los límites de carga multiparte de Amazon S3 en la Guía del usuario de Amazon Simple Storage Service.

  > **Warning**
  >
  > Dado que este sistema de archivos heredado puede crear condiciones de carrera que podrían provocar daños en el sistema de archivos, debe evitar este formato y usar EMRFS en su lugar.

```
s3bfs://bucket-name/path-to-file-in-bucket
```

Preparación de los datos de entrada para su procesamiento con Amazon EMR

La mayoría de los clústeres cargan los datos de entrada y, a continuación, procesan dichos datos. Para cargar datos, tiene que estar en una ubicación a la que el clúster pueda acceder y en un formato que el clúster pueda procesar. El escenario más habitual consiste en cargar los datos de entrada en Amazon S3. Amazon EMR proporciona herramientas para su clúster para importar o leer datos desde Amazon S3.

El formato de entrada predeterminado en Hadoop son archivos de texto, aunque puede personalizar Hadoop y utilizar herramientas para importar los datos almacenados en otros formatos.

Temas

- Tipos de entrada que Amazon EMR puede aceptar
- Diferentes formas de obtener datos en Amazon EMR

Tipos de entrada que Amazon EMR puede aceptar

El formato de entrada predeterminado para un clúster son archivos de texto con cada línea separada por un carácter de nueva línea (\n), que es el formato de entrada usado con más frecuencia.

Si los datos de entrada se encuentran en un formato que no sean los archivos de texto predeterminados, puede utilizar la interfaz de Hadoop `InputFormat` para especificar otros tipos de entrada. Puede incluso crear una subclase de la clase `FileInputFormat` para gestionar tipos de datos personalizados. Para obtener más información, consulte http://hadoop.apache. org/docs/current/api/org/apache/hadoop/mapred/InputFormat.html.

Si usa Hive, puede usar a serializer/deserializer (SerDe) para leer datos de un formato determinado en HDFS. Para obtener más información, consulte https://cwiki.apache. org/confluence/display/Hive/SerDe.

Diferentes formas de obtener datos en Amazon EMR

Amazon EMR proporciona varias formas de obtener datos en un clúster. La forma más común consiste en cargar los datos en Amazon S3 y utilizar las características integradas de Amazon EMR para cargar los datos en el clúster. También puede utilizar la característica DistributedCache de Hadoop para transferir archivos desde un sistema de archivos distribuido al sistema de archivos local. La implementación de Hive proporcionada por Amazon EMR (Hive versión 0.7.1.1 y posteriores) incluye la funcionalidad que puede utilizar para importar y exportar datos entre DynamoDB y un clúster de Amazon EMR. Si tiene una gran cantidad de datos on-premises para procesar, puede encontrar útil el servicio AWS Direct Connect .

Temas

- Descargar datos en Amazon S3
- Carga de datos con AWS DataSync
- Importación de archivos con caché distribuida con Amazon EMR
- Detección y procesamiento de archivos comprimidos con Amazon EMR
- Importación de datos de DynamoDB a Hive con Amazon EMR
- Conexión a los datos con AWS Direct Connect desde Amazon EMR
- Carga de grandes cantidades de datos para Amazon EMR con AWS Snowball Edge

Descargar datos en Amazon S3

Para obtener información sobre cómo cargar objetos en Amazon S3, consulte Agregar un objeto a un bucket en la Guía del usuario de Amazon Simple Storage Service. Para obtener más información sobre el uso de Amazon S3 con Hadoop, consulte http://wiki.apache. org/hadoop/AmazonS3.

Temas

- Creación y configuración de un bucket de Amazon S3
- Configurar cargas multiparte para Amazon S3
- Prácticas recomendadas
- Carga de datos en Amazon S3 Express One Zone

Creación y configuración de un bucket de Amazon S3

Amazon EMR lo utiliza AWS SDK para Java con Amazon S3 para almacenar datos de entrada, archivos de registro y datos de salida. Amazon S3 denomina estas ubicaciones de almacenamiento

buckets. Los buckets tienen ciertas restricciones y limitaciones para cumplir con los requisitos de Amazon S3 y DNS. Para obtener más información, consulte Restricciones y limitaciones de los buckets en la Guía del usuario de Amazon Simple Storage Service.

En esta sección, se muestra cómo utilizar Amazon S3 AWS Management Console para crear y, a continuación, establecer los permisos para un bucket de Amazon S3. También puede crear y establecer permisos para un bucket de Amazon S3 con la API o la AWS CLI de Amazon S3. También puede utilizar curl junto con una modificación para pasar los parámetros de autenticación correspondientes a Amazon S3.

Consulte los siguientes recursos:

- Para crear un bucket mediante la consola, consulte Crear un bucket en la Guía del usuario de Amazon S3.
- Para crear buckets y trabajar con ellos mediante el AWS CLI, consulte Uso de comandos de S3 de alto nivel AWS Command Line Interface en la Guía del usuario de Amazon S3.
- Para crear un bucket mediante un SDK, consulte Ejemplos de creación de un bucket en la Guía del usuario de Amazon Simple Storage Service.
- Para trabajar con buckets mediante curl, consulte Herramienta de autenticación de Amazon S3 para curl.
- Para más información acerca de la especificación de buckets específicos de una región, consulte Acceso a un bucket en la Guía del usuario de Amazon Simple Storage Service.
- Para trabajar con buckets mediante Puntos de acceso de Amazon S3, consulte Uso de un alias de estilo de bucket para su punto de acceso en la Guía del usuario de Amazon S3. Puede utilizar fácilmente Puntos de acceso de Amazon S3 con el alias del punto de acceso de Amazon S3 en lugar del nombre del bucket de Amazon S3. Puede utilizar el alias del punto de acceso Amazon S3 para aplicaciones nuevas y existentes, incluidas Spark, Hive y Presto, entre otras.

> **Note**
>
> Si habilita el registro para un bucket, se habilitan únicamente los registros de acceso al bucket, no los registros del clúster de Amazon EMR.

Durante la creación de un bucket o posteriormente, puede definir los permisos adecuados para tener acceso a él, en función de su aplicación. Lo habitual es que se conceda a sí mismo (el propietario) acceso de lectura y de escritura, así como acceso de lectura a los usuarios autenticados.

Los buckets de Amazon S3 necesarios deben existir para poder crear un clúster. Debe cargar en Amazon S3 todos los scripts o datos necesarios a los que se hace referencia en el clúster.

Configurar cargas multiparte para Amazon S3

Amazon EMR admite la carga multiparte de Amazon S3 a través del AWS SDK for Java. La carga multiparte permite cargar un solo objeto como un conjunto de partes. Puede cargar estas partes del objeto de forma independiente y en cualquier orden. Si la transmisión de cualquier parte falla, puede retransmitir esta parte sin que las demás partes se vean afectadas. Después de cargar todas las partes del objeto, Amazon S3 las combina y crea el objeto.

Para obtener más información, consulte la sección Información general sobre la carga multiparte en la Guía de del usuario de Amazon Simple Storage Service.

Además, Amazon EMR ofrece propiedades que le permiten controlar de manera más precisa la limpieza de partes de carga multiparte con errores.

En la siguiente tabla, se describen las propiedades de configuración de Amazon EMR para la carga multiparte. Puede configurarlas utilizando la clasificación de configuración `core-site`. Para más información, consulte Configuración de aplicaciones en la Guía de versiones de Amazon EMR.

Nombre de parámetro de configuración	Valor predeterminado	Descripción
`fs.s3n.multipart.uploads.enabled`	`true`	Un tipo booleano que indica si se debe habilitar las cargas multiparte. Cuando la vista consistente de EMRFS se ha habilitado, las cargas multiparte se habilitan de forma predeterminada y si se configura este valor en `false`, no se tiene en cuenta.
`fs.s3n.multipart.uploads.split.size`	`134217728`	Especifica el tamaño máximo de una parte, en bytes, para que EMRFS inicie la carga de una parte nueva cuando las cargas multiparte están habilitadas. El valor mínimo es `5242880` (5 MB). Si se especifica un valor menor, se utiliza `5242880`. El valor máximo es `5368709120` (5 GB).

Nombre de parámetro de configuración	Valor predeterminado	Descripción
		Si se especifica un valor mayor, se utiliza `5368709120` . Si se ha desactivado el cifrado del cliente de EMRFS y el confirmador optimizado de Amazon S3 también se ha desactivado, este valor también controla el tamaño máximo que puede alcanzar un archivo de datos para que EMRFS utilice cargas multiparte en lugar de una solicitud `PutObject` para cargar el archivo. Para obtener más información, consulte
`fs.s3n.ssl.enabled`	`true`	Un tipo booleano que indica si se debe utilizar http o https.
`fs.s3.buckets.create.enabled`	`false`	Un tipo booleano que indica si se debe crear un bucket en caso de que no exista. Si se establece en `false`, se provocará una excepción en las operaciones `CreateBucket` .
`fs.s3.multipart.clean.enabled`	`false`	Un tipo booleano que indica si se debe habilitar la limpieza periódica en segundo plano de las cargas multiparte incompletas.
`fs.s3.multipart.clean.age.threshold`	`604800`	Un tipo long que especifica la edad mínima de una carga multiparte, en segundos, antes de que se la incluya en la limpieza. El valor predeterminado es una semana.

Nombre de parámetro de configuración	Valor predeterminado	Descripción
`fs.s3.multipart.clean.jitter.max`	`10000`	Un tipo entero que especifica la cantidad máxima de retraso de fluctuación aleatorio, en segundos, que se agrega al retraso fijo de 15 minutos antes de programar la próxima ronda de limpieza.

Desactivar cargas multiparte

Console

Para deshabilitar las cargas multiparte con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Configuración de software, introduzca la configuración siguiente: `classification=core-site,properties=[fs.s3n.multipart.uploads.enabled=false]`.
4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

CLI

Para inhabilitar la carga multiparte mediante el AWS CLI

Este procedimiento explica cómo deshabilitar la carga multiparte utilizando la AWS CLI. Para deshabilitar la carga multiparte, escriba el comando `create-cluster` con el parámetro `--bootstrap-actions`.

1. Cree un archivo `myConfig.json` con el siguiente contenido y guárdelo en el mismo directorio en el que va a ejecutar el comando:

```
[
```

```
  {
    "Classification": "core-site",
    "Properties": {
      "fs.s3n.multipart.uploads.enabled": "false"
    }
  }
]
```

2. Escriba el siguiente comando y *myKey* sustitúyalo por el nombre de su EC2 key pair.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" \
--release-label emr-7.10.0 --applications Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3 --configurations file://myConfig.json
```

API

Para desactivar la carga multiparte mediante la API

- Para obtener información sobre cómo utilizar las cargas multiparte de Amazon S3 mediante programación, consulte Uso del SDK de AWS para Java para cargas multiparte en la Guía del usuario de Amazon Simple Storage Service.

  Para obtener más información sobre el AWS SDK para Java, consulte AWS SDK for Java.

Prácticas recomendadas

A continuación se indican las recomendaciones para utilizar buckets de Amazon S3 con clústeres de EMR.

Habilitación del control de versiones

El control de versiones es una configuración recomendada para su bucket de Amazon S3. Al habilitar el control de versiones, se asegura de que incluso si se eliminan o se sobrescriben involuntariamente los datos, se puedan recuperar. Para más información, consulte Uso del control de versiones en la Guía del usuario de Amazon Simple Storage Service.

Limpiar las cargas multiparte con errores

Los componentes del clúster EMR utilizan cargas multiparte a través del AWS SDK para Java con Amazon S3 para escribir archivos de registro y enviar datos APIs a Amazon S3 de forma predeterminada. Para más información sobre cómo utilizar Amazon EMR para cambiar las propiedades relacionadas con esta configuración, consulte Configurar cargas multiparte para Amazon S3. En ocasiones, la carga de un archivo de gran tamaño puede dar lugar a una carga multiparte de Amazon S3 incompleta. Cuando una carga multiparte no se puede completar de forma satisfactoria, la carga multiparte en curso sigue ocupando su bucket e incurre en costos de almacenamiento. Le recomendamos las siguientes opciones para evitar un almacenamiento de archivos excesivo:

- En el caso de los buckets que utilice con Amazon EMR, utilice una regla de configuración del ciclo de vida en Amazon S3 que elimine las cargas multiparte incompletas tres días después de la fecha de inicio de la carga. Las reglas de configuración del ciclo de vida le permiten controlar la clase de almacenamiento y el ciclo de vida de los objetos. Para más información, consulte Administración del ciclo de vida de los objetos y Anulación de cargas multiparte incompletas con la política de ciclo de vida de buckets.
- Habilite la característica de limpieza de cargas multiparte de Amazon EMR estableciendo `fs.s3.multipart.clean.enabled` en `true` y ajustando otros parámetros de limpieza. Esta característica es útil con volúmenes altos y a gran escala, así como con clústeres que tienen un tiempo de actividad limitado. En este caso, el parámetro `DaysAfterIntitiation` de una regla de configuración de ciclo de vida puede ser demasiado largo, incluso si se establece en su valor mínimo, provocando picos en el almacenamiento de Amazon S3. La limpieza multiparte de Amazon EMR permite un control más preciso. Para obtener más información, consulte Configurar cargas multiparte para Amazon S3.

Administrar los marcadores de versiones

Se recomienda habilitar una regla de configuración de ciclo de vida en Amazon S3 para eliminar los marcadores de eliminación de objetos vencidos de los buckets con control de versiones que

utilice con Amazon EMR. Al eliminar un objeto de un bucket con control de versiones, se crea un marcador de eliminación. Si todas las versiones anteriores del objeto caducan posteriormente, queda un marcador de eliminación de objeto caducado en el bucket. Aunque no se aplican cargos por los marcadores de eliminación, si borra los marcadores caducados puede mejorar el rendimiento de las solicitudes de LIST. Para más información, consulte Configuración del ciclo de vida de un bucket con el control de versiones en la Guía del usuario de Amazon Simple Storage Service.

Prácticas recomendadas de rendimiento

En función de las cargas de trabajo, los tipos específicos de uso de clústeres de EMR y las aplicaciones en dichos clústeres pueden dar lugar a un número elevado de solicitudes en su bucket. Para más información, consulte Consideraciones sobre la tasa de solicitudes y el rendimiento en la Guía del usuario de Amazon Simple Storage Service.

Carga de datos en Amazon S3 Express One Zone

Descripción general

A partir de la versión 6.15.0 de Amazon EMR, puede utilizar Amazon EMR con Apache Spark junto con la clase de almacenamiento Amazon S3 Express One Zone para mejorar el rendimiento de sus trabajos de Spark. Las versiones 7.2.0 y posteriores de Amazon EMR también son compatibles con HBase Flink y Hive, por lo que también puede beneficiarse de S3 Express One Zone si utiliza estas aplicaciones. S3 Express One Zone es una clase de almacenamiento de S3 para aplicaciones que acceden con frecuencia a los datos con cientos de miles de solicitudes por segundo. En el momento de su lanzamiento, S3 Express One Zone ofrece el almacenamiento de objetos en la nube con la latencia más baja y el rendimiento más alto de Amazon S3.

Requisitos previos

- Permisos de S3 Express One Zone: cuando S3 Express One Zone realiza a una acción como `GET`, `LIST` o `PUT` en un objeto de S3, la clase de almacenamiento llama a `CreateSession` en su nombre. Su política de IAM debe permitir el permiso `s3express:CreateSession` para que el conector S3A pueda invocar la API `CreateSession`. Para ver un ejemplo de política con ese permiso, consulte Introducción a Amazon S3 Express One Zone.
- Conector S3A: para configurar el clúster de Spark de modo que pueda acceder a los datos de un bucket de Amazon S3 que utilice la clase de almacenamiento S3 Express One Zone, debe utilizar el conector de Apache Hadoop S3A. Para usar el conector, asegúrese de que todos los S3 URIs utilicen el esquema. `s3a` Si no es así, puede cambiar la implementación del sistema de archivos que utiliza para los esquemas `s3` y `s3n`.

Para cambiar el esquema s3, especifique las siguientes configuraciones de clúster:

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "fs.s3.impl": "org.apache.hadoop.fs.s3a.S3AFileSystem",
      "fs.AbstractFileSystem.s3.impl": "org.apache.hadoop.fs.s3a.S3A"
    }
  }
]
```

Para cambiar el esquema s3n, especifique las siguientes configuraciones de clúster:

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "fs.s3n.impl": "org.apache.hadoop.fs.s3a.S3AFileSystem",
      "fs.AbstractFileSystem.s3n.impl": "org.apache.hadoop.fs.s3a.S3A"
    }
  }
]
```

Introducción a Amazon S3 Express One Zone

Temas

- Creación de una política de permisos
- Creación y configuración de un clúster
- Información general sobre las configuraciones

Creación de una política de permisos

Antes de poder crear un clúster que utilice Amazon S3 Express One Zone, debe crear una política de IAM para adjuntarla al perfil de EC2 instancia de Amazon del clúster. La política debe tener permisos para acceder a la clase de almacenamiento S3 Express One Zone. En la siguiente política de ejemplo, se muestra cómo conceder los permisos necesarios. Después de crear la política, adjúntela al rol de perfil de instancia que utilice para crear su clúster de EMR, tal y como se describe en la sección Creación y configuración de un clúster.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3express:*:123456789012:bucket/example-s3-bucket"
      ],
      "Action": [
        "s3express:CreateSession"
      ],
      "Sid": "AllowS3EXPRESSCreatesession"
    }
  ]
}
```

Creación y configuración de un clúster

A continuación, cree un clúster que ejecute Spark HBase, Flink o Hive con S3 Express One Zone. En los siguientes pasos, se ofrece información general de alto nivel para crear un clúster en la AWS Management Console:

1. Navegue hasta la consola de Amazon EMR y seleccione Clústeres en la barra lateral. A continuación, elija Crear clúster.
2. Si utiliza Spark, seleccione la versión `emr-6.15.0` o posterior de Amazon EMR. Si usas HBase Flink o Hive, selecciona o una opción superior. `emr-7.2.0`
3. Seleccione las aplicaciones que desee incluir en el clúster, como Spark o HBase Flink.
4. Para habilitar Amazon S3 Express One Zone, introduzca una configuración similar a la del siguiente ejemplo en la sección Configuración de software. Las configuraciones y los valores recomendados se describen en la sección Información general sobre las configuraciones que sigue a este procedimiento.

```
[
  {
    "Classification": "core-site",
    "Properties": {
```

```
      "fs.s3a.aws.credentials.provider":
 "software.amazon.awssdk.auth.credentials.InstanceProfileCredentialsProvider",
      "fs.s3a.change.detection.mode": "none",
      "fs.s3a.endpoint.region": "aa-example-1",
      "fs.s3a.select.enabled": "false"
    }
  },
  {
    "Classification": "spark-defaults",
    "Properties": {
      "spark.sql.sources.fastS3PartitionDiscovery.enabled": "false"
    }
  }
]
```

5. En la sección Perfil de EC2 instancia para Amazon EMR, elija usar un rol existente y use un rol con la política adjunta que creó en la Creación de una política de permisos sección anterior.
6. Configure el resto de los ajustes del clúster según corresponda para su aplicación y, a continuación, seleccione Crear clúster.

Información general sobre las configuraciones

En las siguientes tablas, se describen las configuraciones y los valores sugeridos que debe especificar al configurar un clúster que utiliza S3 Express One Zone con Amazon EMR, tal y como se describe en la sección Creación y configuración de un clúster.

Configuraciones de S3A

Parámetro	Valor predeterminado	Valor sugerido	Explicación
`fs.s3a.aws.credentials.provider`	Si no se especifica, se utiliza `AWSCredentialProviderList` en el siguiente orden: `TemporaryAWSCredentialsProvider`, `SimpleAWS`	`software.amazon.awssdk.auth.credentials.InstanceProfileCredentialsProvider`	El rol del perfil de instancia de Amazon EMR debe tener la política que permita al sistema de archivos S3A llamar a `s3express:CreateSession`. Otros proveedores de

Parámetro	Valor predeterminado	Valor sugerido	Explicación
	`CredentialsProvider`, `EnvironmentVariableCredentialsProvider`, `IAMInstanceCredentialsProvider`.		credenciales también funcionan si tienen los permisos de S3 Express One Zone.
`fs.s3a.endpoint.region`	null	El Región de AWS lugar en el que creó el bucket.	La lógica de resolución regional no funciona con la clase de almacenamiento S3 Express One Zone.
`fs.s3a.select.enabled`	`true`	`false`	`select` de Amazon S3 no es compatible con la clase de almacenamiento S3 Express One Zone.
`fs.s3a.change.detection.mode`	`server`	none	La detección de cambios de S3A se realiza al comprobar las `etags` basadas en MD5. La clase de almacenamiento S3 Express One Zone no es compatible con `checksums` de MD5.

Configuraciones de Spark

Parámetro	Valor predeterminado	Valor sugerido	Explicación
`spark.sql.sources.fastS3PartitionDiscovery.enabled`	`true`	`false`	La optimización interna utiliza un parámetro de API de S3 que la clase de almacenamiento S3 Express One Zone no admite.

Configuraciones de Hive

Parámetro	Valor predeterminado	Valor sugerido	Explicación
`hive.exec.fast.s3.partition.discovery.enabled`	`true`	`false`	La optimización interna utiliza un parámetro de API de S3 que la clase de almacenamiento S3 Express One Zone no admite.

Consideraciones

Tenga en cuenta lo siguiente al integrar Apache Spark en Amazon EMR con la clase de almacenamiento S3 Express One Zone:

- Se requiere el conector S3A para utilizar S3 Express One Zone con Amazon EMR. Solo S3A tiene las características y las clases de almacenamiento necesarias para interactuar con S3 Express One Zone. Para ver los pasos para configurar el conector, consulte the section called “Requisitos previos”.
- La clase de almacenamiento Amazon S3 Express One Zone admite SSE-S3 SSE-KMS cifrado. Para obtener más información, consulte Cifrado del lado del servidor con Amazon S3.

- La clase de almacenamiento Amazon S3 Express One Zone no es compatible con la escritura con el `FileOutputCommitter` de S3A. Si se escribe con el `FileOutputCommitter` de S3A en los buckets de S3 Express One Zone, se produce un error: InvalidStorageClass: The storage class you specified is not valid.
- Amazon S3 Express One Zone es compatible con las versiones 6.15.0 y posteriores de Amazon EMR a partir de EMR. EC2 Además, es compatible con las versiones 7.2.0 y posteriores de Amazon EMR en Amazon EMR con EKS y en Amazon EMR sin servidor.

Carga de datos con AWS DataSync

AWS DataSync es un servicio de transferencia de datos en línea que simplifica, automatiza y acelera el proceso de transferencia de datos entre el almacenamiento local y los servicios de almacenamiento o entre los servicios de AWS almacenamiento. AWS DataSync admite diversos sistemas de almacenamiento local, como el sistema de archivos distribuido Hadoop (HDFS), los servidores de archivos NAS y el almacenamiento de objetos autogestionado.

La forma más común de tener datos en un clúster es mediante la carga de datos en Amazon S3 y el uso de características integradas de Amazon EMR para cargar los datos en el clúster.

DataSync puede ayudarle a realizar las siguientes tareas:

- Replicar el HDFS de su clúster de Hadoop en Amazon S3 para garantizar la continuidad empresarial
- Copiar el HDFS a Amazon S3 para rellenar sus lagos de datos
- Transferir datos entre el HDFS de su clúster de Hadoop y Amazon S3 para su análisis y procesamiento

Para cargar datos a su depósito de S3, primero debe implementar uno o más DataSync agentes en la misma red que su almacenamiento local. Un agente es una máquina virtual (VM) que se utiliza para leer o escribir datos en una ubicación autoadministrada. A continuación, activa los agentes en el depósito de S3 Cuenta de AWS y en el Región de AWS lugar en el que se encuentra.

Una vez activado el agente, debe crear una ubicación de origen para el almacenamiento en las instalaciones, una ubicación de destino para el bucket de S3 y una tarea. Una tarea es un conjunto de dos ubicaciones (origen y destino) y un conjunto de opciones predeterminadas que se utilizan para controlar el comportamiento de la tarea.

Por último, ejecuta la DataSync tarea de transferir los datos del origen al destino.

Para obtener más información, consulta Introducción a AWS DataSync.

Importación de archivos con caché distribuida con Amazon EMR

DistributedCache es una característica de Hadoop que puede aumentar la eficacia cuando una tarea de asignación o de reducción necesita acceder a datos comunes. Si el clúster depende de aplicaciones existentes o binarias que no se instalaron al crear el clúster, puede utilizar DistributedCache para importar estos archivos. Esta característica permite a un nodo de clúster leer los archivos importados desde su sistema de archivos local, en lugar de recuperar los archivos desde otros nodos del clúster.

Para obtener más información, vaya a http://hadoop.apache. org/docs/stable/api/org/apache/hadoop/filecache/DistributedCache.html.

Invoca DistributedCache al crear el clúster. Los archivos se almacenan en caché justo antes de empezar el trabajo de Hadoop y los archivos permanecen almacenados en la memoria caché durante el tiempo que dura el trabajo. Puede almacenar en caché los archivos almacenados en cualquier sistema de archivos compatible con Hadoop, por ejemplo, HDFS o Amazon S3. El tamaño predeterminado de la caché de archivos es de 10 GB. Para cambiar el tamaño de la caché, vuelva a configurar el parámetro de Hadoop, `local.cache.size` a través de la acción de arranque. Para obtener más información, consulte Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR.

Temas

- Tipos de archivo admitidos
- Ubicación de los archivos en caché
- Acceso a archivos almacenados en caché desde aplicaciones de streaming
- Acceso a archivos almacenados en caché desde aplicaciones de streaming

Tipos de archivo admitidos

DistributedCache permite importar tanto archivos individuales como conjuntos de archivos. Los archivos individuales se almacenan en caché como de solo lectura. Los archivos ejecutables y binarios tienen establecidos permisos de ejecución.

Los conjuntos de archivos son uno o más archivos empaquetados mediante una utilidad, como `gzip`. DistributedCache transfiere los archivos comprimidos a cada nodo central y descomprime el

conjunto de archivos como parte del proceso de almacenamiento en caché. DistributedCache es compatible con los siguientes formatos de compresión:

- zip
- tgz
- tar.gz
- tar
- jar

Ubicación de los archivos en caché

DistributedCache solo copia los archivos en los nodos centrales. Si no hay nodos centrales en el clúster, DistributedCache copia los archivos en el nodo principal.

DistributedCache asocia los archivos en caché al directorio de trabajo actual del asignador y reductor mediante enlaces simbólicos. Un symlink es un alias a una ubicación de archivo, no la ubicación de archivo real. El valor del parámetro, `yarn.nodemanager.local-dirs` en `yarn-site.xml`, especifica la ubicación de los archivos temporales. Amazon EMR establece este parámetro en `/mnt/mapred` o alguna variación basada en el tipo de instancia y la versión de EMR. Por ejemplo, una configuración podría tener `/mnt/mapred` y `/mnt1/mapred` dado que el tipo de instancia tiene dos volúmenes efímeros. Los archivos de caché se encuentran en un subdirectorio de la ubicación de archivos temporales en `/mnt/mapred/taskTracker/archive`.

Si almacena un solo archivo en caché, DistributedCache colocará el archivo en el directorio `archive`. Si almacena un conjunto de archivos en caché, DistributedCache lo descomprime y crea un subdirectorio en `/archive` con el mismo nombre que el del conjunto de archivos. Los archivos individuales se encuentran en el nuevo subdirectorio.

Puede utilizar DistributedCache solo cuando se utilice una transmisión.

Acceso a archivos almacenados en caché desde aplicaciones de streaming

Para acceder a los archivos almacenados en la memoria caché de sus aplicaciones de mapeador o reductor, asegúrese de añadir el directorio de trabajo actual (./) en la ruta de la aplicación y hacer referencia a los archivos en caché aunque estén presentes en el directorio de trabajo actual.

Acceso a archivos almacenados en caché desde aplicaciones de streaming

Puede usar el AWS Management Console y el AWS CLI para crear clústeres que usen caché distribuida.

Console

Para especificar los archivos de caché distribuida utilizando la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Pasos, seleccione Agregar paso. Se abrirá el cuadro de diálogo Agregar paso. En el campo Argumentos, incluya los archivos y conjuntos de archivos que desea guardar en la memoria caché. El tamaño del archivo (o tamaño total de los archivos en un archivo de almacenamiento) debe ser menor que el tamaño de la memoria caché asignado.

   Si desea agregar un archivo individual en la caché distribuida, especifique `-cacheFile` seguido del nombre y ubicación del archivo, el símbolo numeral (#) y el nombre que desee asignarle al archivo cuando se coloque en la caché local. En el ejemplo siguiente se muestra cómo agregar un archivo individual a la caché distribuida.

   ```
   -cacheFile \
   s3://amzn-s3-demo-bucket/file-name#cache-file-name
   ```

   Si desea agregar un conjunto de archivos en la caché distribuida, introduzca `-cacheArchive` seguido de la ubicación de los archivos en Amazon S3, el símbolo numeral (#) y a continuación el nombre que desee asignarle a la colección de archivos en la caché local. En el ejemplo siguiente se muestra cómo agregar un archivo de almacenamiento a la caché distribuida.

   ```
   -cacheArchive \
   s3://amzn-s3-demo-bucket/archive-name#cache-archive-name
   ```

   Ingrese los valores adecuados en el resto de los campos del cuadro de diálogo. Las opciones varían según el tipo de paso. Para agregar el paso y salir del cuadro de diálogo, elija Agregar paso.

4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

CLI

Para especificar los archivos de caché distribuidos con el AWS CLI

- Para enviar un paso de streaming cuando se crea un clúster, escriba el comando `create-cluster` con el parámetro `--steps`. Para especificar los archivos de caché distribuidos mediante el AWS CLI, especifique los argumentos adecuados al enviar un paso de transmisión.

  Si desea agregar un archivo individual en la caché distribuida, especifique `-cacheFile` seguido del nombre y ubicación del archivo, el símbolo numeral (#) y el nombre que desee asignarle al archivo cuando se coloque en la caché local.

  Si desea agregar un conjunto de archivos en la caché distribuida, introduzca `-cacheArchive` seguido de la ubicación de los archivos en Amazon S3, el símbolo numeral (#) y a continuación el nombre que desee asignarle a la colección de archivos en la caché local. En el ejemplo siguiente se muestra cómo agregar un archivo de almacenamiento a la caché distribuida.

  Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Example 1

Escriba el siguiente comando para lanzar un clúster y enviar un paso de streaming que utiliza `-cacheFile` para añadir un archivo `sample_dataset_cached.dat`, a la caché.

```
aws emr create-cluster --name "Test cluster" --release-label emr-4.0.0 --
applications 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","s3://my_bucket/my_mapper.py
 s3://my_bucket/my_reducer.py","-mapper","my_mapper.py","-reducer","my_reducer.py,"-
input","s3://my_bucket/my_input","-output","s3://my_bucket/my_output", "-
cacheFile","s3://my_bucket/sample_dataset.dat#sample_dataset_cached.dat"]
```

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos básicos. Todos los nodos utilizarán el tipo de instancia especificado en el comando.

Si no ha creado anteriormente el rol de servicio EMR y el perfil de EC2 instancia predeterminados, escriba `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Example 2

El siguiente comando muestra la creación de un clúster de streaming y utiliza `-cacheArchive` para añadir un archivo de almacenamiento de archivos a la caché.

```
aws emr create-cluster --name "Test cluster" --release-label emr-4.0.0 --
applications 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","s3://my_bucket/my_mapper.py
 s3://my_bucket/my_reducer.py","-mapper","my_mapper.py","-reducer","my_reducer.py,"-
input","s3://my_bucket/my_input","-output","s3://my_bucket/my_output", "-
cacheArchive","s3://my_bucket/sample_dataset.tgz#sample_dataset_cached"]
```

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos básicos. Todos los nodos utilizarán el tipo de instancia especificado en el comando.

Si no ha creado anteriormente el rol de servicio EMR y el perfil de EC2 instancia predeterminados, escriba `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Detección y procesamiento de archivos comprimidos con Amazon EMR

Hadoop comprueba la extensión de archivo para detectar archivos comprimidos. Los tipos de compresión compatibles con Hadoop son: gzip, bzip2 y LZO. No tiene que tomar ninguna medida adicional para extraer archivos utilizando estos tipos de compresión; Hadoop se encarga por usted.

Para indexar los archivos LZO, puede usar la biblioteca hadoop-lzo, que se puede descargar de hadoop-lzo. https://github.com/kevinweil/ Tenga en cuenta que, dado que se trata de una biblioteca de terceros, Amazon EMR no ofrece soporte al desarrollador sobre cómo utilizar esta herramienta. Para obtener información sobre su uso, consulte el archivo readme de hadoop-lzo.

Importación de datos de DynamoDB a Hive con Amazon EMR

La implementación de Hive proporcionada por Amazon EMR incluye funcionalidad que puede utilizar para importar y exportar datos entre DynamoDB y un clúster de Amazon EMR. Esto resulta útil si los datos de entrada se almacenan en DynamoDB. Para más información, consulte Exportación, importación, consulta y unión de tablas en DynamoDB mediante Amazon EMR.

Conexión a los datos con AWS Direct Connect desde Amazon EMR

AWS Direct Connect es un servicio que puede utilizar para establecer una conexión de red privada dedicada a Amazon Web Services desde su centro de datos, oficina o entorno de colocación. Si tiene grandes cantidades de datos de entrada, su uso AWS Direct Connect puede reducir los costos de la red, aumentar el rendimiento del ancho de banda y proporcionar una experiencia de red más uniforme que las conexiones basadas en Internet. Para obtener más información, consulte la Guía del usuario de AWS Direct Connect.

Carga de grandes cantidades de datos para Amazon EMR con AWS Snowball Edge

AWS Snowball Edge es un servicio que puede utilizar para transferir grandes cantidades de datos entre Amazon Simple Storage Service (Amazon S3) y su ubicación faster-than-internet de almacenamiento de datos in situ a gran velocidad. Snowball Edge admite dos tipos de trabajos: trabajos de importación y trabajos de exportación. Los trabajos de importación implican una transferencia de datos desde un origen en las instalaciones a un bucket de Amazon S3. Los trabajos de exportación implican una transferencia de datos de un bucket de Amazon S3 a un origen en las instalaciones. Para ambos tipos de trabajo, los dispositivos Snowball Edge protegen y protegen sus datos, mientras que las empresas de transporte regionales los transportan entre Amazon S3 y su ubicación de almacenamiento de datos in situ. Los dispositivos Snowball Edge son físicamente robustos y están protegidos por el AWS Key Management Service ()AWS KMS. Para más información, consulte Guía para desarrolladores de AWS Snowball Edge Edge.

Cómo configurar una ubicación para la salida del clúster de Amazon EMR

El formato de salida más común de un clúster de Amazon EMR son los archivos de texto, ya sea comprimidos o sin comprimir. Normalmente, están escritos en un bucket de Amazon S3. Este bucket se debe crear antes de lanzar el clúster. Puede especificar el bucket de S3 como ubicación de salida al lanzar el clúster.

Para obtener más información, consulte los temas siguientes:

Temas

- Creación y configuración de un bucket de Amazon S3
- ¿Qué formatos puede devolver Amazon EMR?
- Cómo escribir datos en un bucket de Amazon S3 que no le pertenece con Amazon EMR
- Formas de comprimir la salida de su clúster de Amazon EMR

Creación y configuración de un bucket de Amazon S3

Amazon EMR (Amazon EMR) utiliza Amazon S3 para almacenar datos de entrada y de salida, y archivos de registro. Amazon S3 denomina estas ubicaciones de almacenamiento buckets. Los buckets tienen ciertas restricciones y limitaciones para cumplir con los requisitos de Amazon S3 y DNS. Para obtener más información, consulte Restricciones y limitaciones de los buckets en la Guía para desarrolladores de Amazon Simple Storage Service.

Para crear un bucket de Amazon S3, siga las instrucciones que se muestran en Creación de un bucket en la Guía del usuario de Amazon Simple Storage Service.

> **Note**
>
> Si habilita el registro en el asistente de Crear un bucket, solo se habilitan los registros de acceso al bucket, no los registros del clúster.

> **Note**
>
> Para obtener más información sobre la especificación de depósitos específicos de una región, consulte Depósitos y regiones en la Guía para desarrolladores de Amazon Simple Storage Service y los puntos de enlace regionales disponibles para. AWS SDKs

Después de crear su bucket, puede definir los permisos correspondientes. Lo habitual es que se conceda (el propietario) acceso de lectura y escritura. Se recomienda encarecidamente que siga las prácticas recomendadas de seguridad para Amazon S3 al configurar su bucket.

Los buckets de Amazon S3 necesarios deben existir para poder crear un clúster. Debe cargar en Amazon S3 todos los scripts o datos necesarios a los que se hace referencia en el clúster. En la siguiente tabla se describen ubicaciones de ejemplo de datos, scripts y archivos de registro.

Información	Ejemplo de ubicación en Amazon S3
script o programa	`s3://amzn-s3-demo-bucket1/script/MapperScript.py`
archivos de registro	`s3://amzn-s3-demo-bucket1/logs`
datos de entrada	`s3://amzn-s3-demo-bucket1/input`
datos de salida	`s3://amzn-s3-demo-bucket1/output`

¿Qué formatos puede devolver Amazon EMR?

El formato de salida predeterminado para un clúster es texto con parejas de clave, valor escritas en líneas individuales de los archivos de texto. Este es el formato de salida usado con más frecuencia.

Si los datos de salida se tienen que escribir en un formato que no sean los archivos de texto predeterminados, puede utilizar la interfaz de Hadoop `OutputFormat` para especificar otros tipos de salida. Puede incluso crear una subclase de la clase `FileOutputFormat` para gestionar tipos de datos personalizados. Para obtener más información, consulte http://hadoop.apache. org/docs/current/api/org/apache/hadoop/mapred/OutputFormat.html.

Si va a lanzar un clúster de Hive, puede usar a serializer/deserializer (SerDe) para enviar datos de HDFS a un formato determinado. Para obtener más información, consulte https://cwiki.apache. org/confluence/display/Hive/SerDe.

Cómo escribir datos en un bucket de Amazon S3 que no le pertenece con Amazon EMR

Cuando escriba un archivo en un bucket de Amazon Simple Storage Service (Amazon S3), de forma predeterminada, usted es el único capaz de leer dicho archivo. La suposición es que escribirá archivos en sus propios buckets y esta configuración predeterminada protege la privacidad de sus archivos.

Sin embargo, si está ejecutando un clúster y quiere que el resultado se escriba en el bucket de Amazon S3 de otro AWS usuario y quiere que ese otro AWS usuario pueda leer ese resultado, debe hacer dos cosas:

- Pide al otro AWS usuario que te conceda permisos de escritura para su bucket de Amazon S3. El clúster que lance se ejecuta con sus AWS credenciales, por lo que cualquier clúster que lance también podrá escribir en el bucket de ese otro AWS usuario.
- Establezca permisos de lectura para el otro AWS usuario en los archivos que usted o el clúster escriban en el bucket de Amazon S3. La forma más sencilla de configurar estos permisos de lectura es utilizar listas de control de acceso predefinidas (ACLs), un conjunto de políticas de acceso predefinidas definidas por Amazon S3.

Para obtener información sobre cómo el otro AWS usuario puede concederle permisos para escribir archivos en el bucket de Amazon S3 del otro usuario, consulte Edición de permisos de bucket en la Guía del usuario de Amazon Simple Storage Service.

Para que su clúster utilice archivos predefinidos ACLs cuando escriba archivos en Amazon S3, defina la opción de configuración del `fs.s3.canned.acl` clúster en la ACL predefinida que desee utilizar. En la siguiente tabla se enumeran las listas predefinidas actualmente ACLs.

ACL predefinidas	Descripción
`AuthenticatedRead`	Especifica que al usuario se le otorga `Permission.FullControl` y al beneficiario del grupo `GroupGrantee.AuthenticatedUsers` se le otorga acceso `Permission.Read` .
`BucketOwnerFullControl`	Especifica que al propietario del bucket se le otorga `Permission.FullControl` . El propietario del bucket no es necesariamente el mismo que el propietario del objeto.
`BucketOwnerRead`	Especifica que al propietario del bucket se le otorga `Permission.Read` . El propietario del bucket no es necesariamente el mismo que el propietario del objeto.
`LogDeliveryWrite`	Especifica que al propietario se le otorga `Permission.FullControl` y al beneficiario del grupo `GroupGrantee.LogDelivery` se le otorga acceso `Permission.Write` , de modo que se puedan enviar los registros de acceso.

ACL predefinidas	Descripción
`Private`	Especifica que al propietario se le otorga `Permission.FullControl` .
`PublicRead`	Especifica que al usuario se le otorga `Permission.FullControl` y al beneficiario del grupo `GroupGrantee.AllUsers` se le otorga acceso `Permission.Read` .
`PublicReadWrite`	Especifica que al usuario se le otorga `Permission.FullControl` y al beneficiario del grupo `GroupGrantee.AllUsers` se le otorga acceso `Permission.Read` y `Permission.Write` .

Existen muchas formas de definir las opciones de configuración del clúster, en función del tipo de clúster que esté ejecutando. Los siguientes procedimientos muestran cómo definir la opción para casos comunes.

Para escribir archivos usando archivos predefinidos ACLs en Hive

- Desde el mensaje de comando de Hive, establezca la opción de configuración `fs.s3.canned.acl` en la ACL predefinida que desea que el clúster defina para los archivos que escriba en Amazon S3. Para acceder a la línea de comandos de Hive conéctese al nodo principal mediante SSH y escriba Hive en la línea de comandos de Hadoop. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.

  El ejemplo siguiente establece la opción de configuración `fs.s3.canned.acl` en `BucketOwnerFullControl`, que proporciona al propietario del bucket de Amazon S3 un control completo sobre el archivo. Tenga en cuenta que el comando definido distingue entre mayúsculas y minúsculas y no contiene comillas ni espacios.

  ```
  hive> set fs.s3.canned.acl=BucketOwnerFullControl;
  create table acl (n int) location 's3://amzn-s3-demo-bucket/acl/';
  insert overwrite table acl select count(*) from acl;
  ```

Las dos últimas líneas del ejemplo crean una tabla que se almacena en Amazon S3 y escribe datos en la tabla.

Para escribir archivos utilizando la versión preconfigurada de ACLs Pig

- Desde la línea de comandos de Pig, establezca la opción de configuración `fs.s3.canned.acl` en la ACL predefinida que desea que el clúster defina para los archivos que escriba en Amazon S3. Para acceder a la línea de comandos de Pig conéctese al nodo principal mediante SSH y escriba Pig en la línea de comandos de Hadoop. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.

  En el siguiente ejemplo, se establece la opción de `fs.s3.canned.acl` configuración en BucketOwnerFullControl, lo que proporciona al propietario del bucket de Amazon S3 el control total sobre el archivo. Tenga en cuenta que el comando definido incluye un espacio antes del nombre de ACL predefinido y no contiene comillas.

```
pig> set fs.s3.canned.acl BucketOwnerFullControl;
store some data into 's3://amzn-s3-demo-bucket/pig/acl';
```

Para escribir archivos utilizando archivos ACLs predefinidos en un JAR personalizado

- Defina la opción de configuración `fs.s3.canned.acl` utilizando Hadoop con la marca -D. Esto se muestra en el siguiente ejemplo.

```
hadoop jar hadoop-examples.jar wordcount
-Dfs.s3.canned.acl=BucketOwnerFullControl s3://amzn-s3-demo-bucket/input s3://amzn-
s3-demo-bucket/output
```

Formas de comprimir la salida de su clúster de Amazon EMR

Existen diferentes maneras de comprimir la salida que resulta del procesamiento de datos. Las herramientas de compresión que utilice dependen de las propiedades de los datos. La compresión puede mejorar el rendimiento al transferir grandes cantidades de datos.

Compresión de datos de salida

Esto comprime la salida del trabajo de Hadoop. Si lo está utilizando, TextOutputFormat el resultado es un archivo de texto comprimido con gzip. Si está escribiendo SequenceFiles , el resultado es un archivo comprimido SequenceFile internamente. Esto se puede habilitar definiendo el ajuste de configuración mapred.output.compress en true.

Si está ejecutando un trabajo de streaming puede habilitarlo transfiriendo estos argumentos al trabajo de streaming.

```
-jobconf mapred.output.compress=true
```

También puede utilizar una acción de arranque para comprimir automáticamente todas las salidas de trabajo. Aquí se muestra cómo hacerlo con el cliente de Ruby.

```
--bootstrap-actions s3://elasticmapreduce/bootstrap-actions/configure-hadoop \
--args "-s,mapred.output.compress=true"
```

Por último, si se escribe un JAR personalizado puede habilitar la compresión de salida con la siguiente línea al crear el trabajo.

```
FileOutputFormat.setCompressOutput(conf, true);
```

Compresión de datos intermedia

Si el trabajo reorganiza una cantidad de datos importante desde los mapeadores a los reductores, puede obtener una mejora del rendimiento habilitando la compresión intermedia. Comprima la salida de mapeado y descomprímala cuando llegue al nodo secundario. El ajuste de configuración es mapred.compress.map.output. Puede habilitarlo de manera similar a la compresión de salida.

Al escribir un archivo JAR personalizado, utilice el siguiente comando:

```
conf.setCompressMapOutput(true);
```

Uso de la biblioteca Snappy con Amazon EMR

Snappy es una biblioteca de compresión y descompresión optimizada para velocidad. Está disponible en la AMIs versión 2.0 y posteriores de Amazon EMR y se usa de forma predeterminada para la compresión intermedia. Para obtener más información sobre Snappy, consulte http://code.google.com/p/snappy/.

Planificación y configuración de nodos principales en su clúster de Amazon EMR

Cuando se lanza un clúster de Amazon EMR, es posible tener uno o tres nodos principales en el clúster. La alta disponibilidad para flotas de instancias es compatible a partir de las versiones 5.36.1, 5.36.2, 6.8.1, 6.9.1, 6.10.1, 6.11.1 y 6.12.0 y posteriores de Amazon EMR. Para grupos de instancias, la alta disponibilidad es compatible a partir de la versión 5.23.0 de Amazon EMR. Para mejorar aún más la disponibilidad de los clústeres, Amazon EMR puede utilizar los grupos de EC2 ubicación de Amazon para garantizar que los nodos principales se coloquen en un hardware subyacente distinto. Para obtener más información, consulte Integración de Amazon EMR con grupos de ubicación EC2 .

Un clúster de Amazon EMR con varios nodos principales ofrece las siguientes ventajas:

- El nodo principal ya no es un punto de error único. Si uno de los nodos principales deja de funcionar, el clúster utiliza los otros dos nodos principales y se ejecuta sin interrupción. Mientras tanto, Amazon EMR sustituye automáticamente el nodo principal que ha dejado de funcionar por uno nuevo que tiene la misma configuración y las mismas acciones de arranque.
- Amazon EMR habilita las funciones de alta disponibilidad de Hadoop de HDFS NameNode y YARN ResourceManager y admite la alta disponibilidad para algunas otras aplicaciones de código abierto.

  Para más información sobre cómo un clúster de Amazon EMR con varios nodos principales es compatible con aplicaciones de código abierto y otras características de Amazon EMR, consulte Características que admiten la alta disponibilidad en un clúster de Amazon EMR y cómo funcionan con aplicaciones de código abierto.

> **Note**
>
> El clúster pueden encontrarse únicamente en una zona de disponibilidad o subred.

Esta sección proporciona información sobre las aplicaciones y características compatibles de un clúster de Amazon EMR, así como los detalles de configuración, prácticas recomendadas y consideraciones para lanzar el clúster.

Temas

- Características que admiten la alta disponibilidad en un clúster de Amazon EMR y cómo funcionan con aplicaciones de código abierto
- Lanzar un clúster de Amazon EMR con varios nodos principales
- Integración de Amazon EMR con grupos de ubicación EC2
- Consideraciones y prácticas recomendadas al crear un clúster de Amazon EMR con varios nodos principales

Características que admiten la alta disponibilidad en un clúster de Amazon EMR y cómo funcionan con aplicaciones de código abierto

En este tema se proporciona información sobre las características de alta disponibilidad de Hadoop de HDFS NameNode y YARN en un clúster de ResourceManager Amazon EMR, y sobre cómo funcionan las características de alta disponibilidad con aplicaciones de código abierto y otras funciones de Amazon EMR.

Alta disponibilidad de HDFS

Un clúster de Amazon EMR con varios nodos principales habilita la característica de alta disponibilidad de HDFS NameNode en Hadoop. Para más información, consulte Alta disponibilidad del HDFS.

En un clúster de Amazon EMR, dos o más nodos independientes se configuran como. NameNodes Uno NameNode está en un `active` estado y los demás están en un `standby` estado. Si el nodo con `active` NameNode errores, Amazon EMR inicia un proceso automático de conmutación por error de HDFS. Un nodo que `standby` NameNode pasa a ser el responsable de todas las operaciones del cliente en el clúster `active` y se hace cargo de ellas. Amazon EMR sustituye el nodo que ha dejado de funcionar por uno nuevo que se une al clúster en el estado `standby`.

> **Note**
>
> En las versiones 5.23.0 a 5.36.2 de Amazon EMR, solo dos de los tres nodos principales ejecutan HDFS. NameNode
> En las versiones 6.x y posteriores de Amazon EMR, los tres nodos principales ejecutan HDFS. NameNode

Si necesita averiguar cuál NameNode es`active`, puede usar SSH para conectarse a cualquier nodo principal del clúster y ejecutar el siguiente comando:

```
hdfs haadmin -getAllServiceState
```

El resultado muestra los nodos en los que NameNode está instalado y su estado. Por ejemplo:

```
ip-##-#-#-##1.ec2.internal:8020 active
ip-##-#-#-##2.ec2.internal:8020 standby
ip-##-#-#-##3.ec2.internal:8020 standby
```

YARN de alta disponibilidad ResourceManager

Un clúster de Amazon EMR con varios nodos principales habilita la función de ResourceManager alta disponibilidad de YARN en Hadoop. Para obtener más información, consulte ResourceManager Alta disponibilidad.

En un clúster de Amazon EMR con varios nodos principales, YARN ResourceManager se ejecuta en los tres nodos principales. Uno ResourceManager está en el `active` estado y los otros dos están en el `standby` estado. Si se produce un `active` ResourceManager error en el nodo principal, Amazon EMR inicia un proceso de conmutación por error automático. Un nodo principal con a `standby` ResourceManager se encarga de todas las operaciones. Amazon EMR reemplaza el nodo principal fallido por uno nuevo, que luego vuelve a unirse al ResourceManager quórum como un. `standby`

Puede conectarse a «http: //:8088/cluster*`master-public-dns-name`*» desde cualquier nodo principal, lo que lo dirigirá automáticamente al administrador de recursos. `active` Para saber qué administrador de recursos tiene el estado `active`, utilice SSH para conectarse a cualquier nodo principal del clúster. A continuación, ejecute el comando siguiente para obtener una lista de los tres nodos principales y su estado:

```
yarn rmadmin -getAllServiceState
```

Aplicaciones admitidas en un clúster de Amazon EMR con varios nodos principales

Puede instalar y ejecutar las siguientes aplicaciones en un clúster de Amazon EMR con varios nodos principales. Para cada aplicación, el proceso de conmutación por error del nodo principal varía.

Aplicación	Disponibilidad durante la conmutación por error del nodo principal	Notas
Flink	La conmutación por error del nodo principal no afecta a la disponibilidad	Los trabajos de Flink en Amazon EMR se ejecutan como aplicaciones YARN. Flink se JobManagers ejecuta como YARN en los nodos principales. ApplicationMasters No JobManager se ve afectado por el proceso de conmutación por error del nodo principal. Si utiliza Amazon EMR versión 5.27.0 o anterior, JobManager se trata de un único punto de error. Cuando se produce un JobManager error, pierde todos los estados de las tareas y no reanudará las tareas en ejecución. Puede habilitar la JobManager alta disponibilidad configurando el recuento de intentos de aplicación, los puntos de control y habilitándolos ZooKeeper como almacenam iento de estado para Flink. Para más informaci ón, consulte Configuración de Flink en un clúster de Amazon EMR con varios nodos principales. A partir de la versión 5.28.0 de Amazon EMR, no es necesaria ninguna configuración manual para habilitar la alta disponibilidad. JobManager

Aplicación	Disponibilidad durante la conmutación por error del nodo principal	Notas
Ganglia	La conmutación por error del nodo principal no afecta a la disponibilidad	Ganglia está disponible en todos los nodos principales, por lo que se puede seguir ejecutando durante el proceso de conmutación por error del nodo principal.
Hadoop	Alta disponibilidad	HDFS NameNode y YARN ResourceManager se conmutan automáticamente por error al nodo en espera cuando se produce un error en el nodo principal activo.
HBase	Alta disponibilidad	HBase cuando se produce un error en el nodo principal activo, se realiza una conmutación automática por error al nodo en espera. Si se conecta a HBase través de un servidor REST o Thrift, debe cambiar a un nodo principal diferente cuando se produzca un error en el nodo principal activo.
HCatalog	La conmutación por error del nodo principal no afecta a la disponibilidad	HCatalog se basa en el metaalmacén de Hive, que existe fuera del clúster. HCatalog permanece disponible durante el proceso de conmutación por error del nodo principal.
JupyterHub	Alta disponibilidad	JupyterHub está instalado en las tres instancia s principales. Se recomienda configurar la persistencia de los cuadernos con el fin de evitar su pérdida en caso de error del nodo principal. Para más información, consulte Configuración de la persistencia de los cuadernos en Amazon S3.

Aplicación	Disponibilidad durante la conmutación por error del nodo principal	Notas
Livy	Alta disponibilidad	Livy se ha instalado en los tres nodos principal es. Cuando deja de funcionar el nodo principal activo, se pierde el acceso a la sesión actual de Livy y es necesario crear una nueva sesión de Livy en otro nodo principal o en el nodo de sustitución nuevo.
Mahout	La conmutación por error del nodo principal no afecta a la disponibilidad	Dado que Mahout no tiene daemon, no se ve afectado por el proceso de conmutación por error del nodo principal.
MXNet	La conmutación por error del nodo principal no afecta a la disponibilidad	Como no MXNet tiene ningún daemon, no se ve afectado por el proceso de conmutación por error del nodo principal.
Phoenix	Alta disponibilidad	Phoenix' solo QueryServer se ejecuta en uno de los tres nodos principales. El Phoenix de los tres maestros está configurado para conectar el Phoenix. QueryServer Puede encontrar la IP privada del servidor de consultas de Phoenix mediante el archivo `/etc/phoenix/conf/phoenix-env.sh`
Pig	La conmutación por error del nodo principal no afecta a la disponibilidad	Dado que Pig no tiene daemon, no se ve afectado por el proceso de conmutación por error del nodo principal.

Aplicación	Disponibilidad durante la conmutación por error del nodo principal	Notas
Spark	Alta disponibilidad	Todas las aplicaciones de Spark se ejecutan en contenedores de YARN y pueden reacciona r ante una conmutación por error del nodo principal de la misma forma que las caracterí sticas de alta disponibilidad de YARN.
Sqoop	Alta disponibilidad	De forma predeterminada, sqoop-job y sqoop-metastore almacenan los datos (descripciones de trabajos) en el disco local del nodo principal que ejecuta el comando. Si desea guardar los datos del metaalmacén en una base de datos externa, consulte la documentación de Apache Sqoop.
Tez	Alta disponibilidad	Dado que los contenedores de Tez se ejecutan en YARN, Tez se comporta de la misma forma que YARN durante el proceso de conmutación por error del nodo principal.
TensorFlow	La conmutación por error del nodo principal no afecta a la disponibilidad	Como no TensorFlow tiene ningún daemon, no se ve afectado por el proceso de conmutación por error del nodo principal.

Aplicación	Disponibilidad durante la conmutación por error del nodo principal	Notas
Zeppelin	Alta disponibilidad	Zeppelin se ha instalado en los tres nodos principales. Zeppelin almacena las notas y las configuraciones de intérprete en HDFS de forma predeterminada para evitar que se pierdan datos. Las sesiones de intérpret e están completamente aisladas en las tres instancias principales. Los datos de la sesión se perderán en caso de error de la instancia principal. Se recomienda no modificar la misma nota simultáneamente en diferentes instancias principales.
ZooKeeper	Alta disponibilidad	ZooKeeper es la base de la función de conmutación por error automática del HDFS. ZooKeeper proporciona un servicio de alta disponibilidad para mantener los datos de coordinación, notificar a los clientes los cambios en esos datos y supervisar los clientes en busca de fallos. Para más información, consulte Conmutación por error automática del HDFS.

Para ejecutar las siguientes aplicaciones en un clúster de Amazon EMR con varios nodos principales, debe configurar una base de datos externa. La base de datos externa se encuentra fuera del clúster y hace que persistan los datos durante el proceso de conmutación por error del nodo principal. Para las siguientes aplicaciones, los componentes de servicio se recuperarán automáticamente durante el proceso de conmutación por error del nodo principal, pero se pueden producir errores en los trabajos activos, en cuyo caso deberán repetirse.

Aplicación	Disponibilidad durante la conmutación por error del nodo principal	Notas
Hive	Alta disponibilidad únicament e para los componentes de servicio	Se requiere un metaalmacén externo para Hive. Debe ser un metaalmacén externo de MySQL, ya que PostgreSQL no es compatibl e con clústeres multimaestros. Para más información, consulte Configuración de un metaalmacén externo para Hive.
Hue	Alta disponibilidad únicament e para los componentes de servicio	Se requiere una base de datos externa para Hue. Para más información, consulte Uso de Hue con una base de datos remota en Amazon RDS.
Oozie	Alta disponibilidad únicament e para los componentes de servicio	Se requiere una base de datos externa para Oozie. Para más información, consulte Uso de Oozie con una base de datos remota en Amazon RDS. Oozie-server y oozie-client se han instalado en los tres nodos principales. Los oozie-client están configurados para conectarse al oozie-server correcto de forma predeterminada.
PrestoDB o PrestoSQL/ Trino	Alta disponibilidad únicament e para los componentes de servicio	Se requiere un metaalmacén de Hive externo para PrestoDB (PrestoSQL en Amazon EMR 6.1.0-6.3.0 o Trino en Amazon EMR 6.4.0 y versiones posteriores). Puedes usar Presto con el catálogo de datos de AWS Glue o usar una base de datos MySQL externa para Hive. La CLI de Presto se ha instalado en los tres nodos principales a fin de que pueda utilizarl a para acceder al coordinador de Presto

Aplicación	Disponibilidad durante la conmutación por error del nodo principal	Notas
		desde cualquiera de los nodos principales. El coordinador de Presto se ha instalado en un solo nodo principal. Puede encontrar el nombre de DNS del nodo principal en el que se ha instalado el coordinador de Presto llamando a la API `describe-cluster` de Amazon EMR y leyendo el valor devuelto del campo `MasterPublicDnsName` en la respuesta.

> **Note**
>
> Cuando un nodo principal deja de funcionar, la conectividad de bases de datos Java (JDBC) o la conectividad de bases de datos abiertas (ODBC) termina su conexión con el nodo principal. Puede conectarse a cualquiera de los demás nodos principales para continuar su trabajo porque el daemon del metaalmacén de Hive se ejecuta en todos los nodos principales. También puede esperar a que se sustituya el nodo principal que ha dejado de funcionar.

Cómo funcionan las características de Amazon EMR en un clúster con varios nodos principales

Conexión con los nodos principales mediante SSH

Puede conectarse con cualquiera de los tres nodos principales de un clúster de Amazon EMR utilizando SSH de la misma forma que se conecta con un nodo principal único. Para más información, consulte Conectarse al nodo principal mediante SSH.

Si un nodo principal deja de funcionar, la conexión SSH a ese nodo principal finaliza. Para continuar con su trabajo, puede conectarse a uno de los otros dos nodos principales. De manera alternativa, puede tener acceso al nuevo nodo principal después de que Amazon EMR sustituya el que ha dejado de funcionar por uno nuevo.

> **Note**
>
> La dirección IP privada del nodo principal de sustitución es la misma que la del anterior. La dirección IP pública del nodo principal de sustitución puede cambiar. Puede recuperar las nuevas direcciones IP en la consola o mediante el comando `describe-cluster` de la CLI de AWS .
>
> NameNode solo se ejecuta en dos de los nodos principales. Sin embargo, puede ejecutar comandos `hdfs` en la CLI y ejecutar trabajos para tener acceso a HDFS en los tres nodos principales.

Cómo trabajar con pasos en un clúster de Amazon EMR con varios nodos principales

Puede enviar pasos a un clúster de Amazon EMR con varios nodos principales del mismo modo que trabaja con pasos en un clúster de un solo nodo principal. Para más información, consulte Enviar trabajo a un clúster.

A continuación se indican consideraciones para trabajar con los pasos de un clúster de Amazon EMR con varios nodos principales:

- Si un nodo principal deja de funcionar, los pasos que se están ejecutando en el nodo principal se marcan como FAILED. Los datos que se hayan escrito localmente se pierden. Sin embargo, el estado FAILED puede que no refleje el estado real de los pasos.
- Si un paso en ejecución ha iniciado una aplicación de YARN cuando el nodo principal deja de funcionar, el paso puede continuar y finalizar correctamente debido a la conmutación por error automática del nodo principal.
- Se recomienda que compruebe el estado de los pasos consultando la salida de los trabajos. Por ejemplo, los MapReduce trabajos utilizan un `_SUCCESS` archivo para determinar si el trabajo se completa correctamente.
- Se recomienda establecer el ActionOnFailure parámetro en CONTINUE o CANCEL_AND_WAIT, en lugar de en TERMINATE_JOB_FLOW o TERMINATE_CLUSTER.

Protección automática contra la terminación

Amazon EMR habilita automáticamente la protección contra la terminación para todos los clústeres con varios nodos principales y anula cualquier configuración de ejecución de pasos que proporcione al crear el clúster. Puede deshabilitar la protección contra la terminación después de que se haya

lanzado el clúster. Consulte Configuración de la protección de terminación para ejecutar clústeres. Para cerrar un clúster con varios nodos principales, primero debe modificar los atributos del clúster para deshabilitar la protección contra la terminación. Para obtener instrucciones, consulte Terminar un clúster de Amazon EMR con varios nodos principales.

Para más información sobre la protección de terminación, consulte Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental.

Características no admitidas en un clúster de Amazon EMR con varios nodos principales

Las siguientes características de Amazon EMR no están disponibles actualmente en un clúster de Amazon EMR con varios nodos principales:

- EMR Notebooks
- Acceso de un clic al servidor del historial de Spark persistente
- Interfaces de usuario de aplicaciones persistentes
- El acceso con un solo clic a las interfaces de usuario de aplicaciones persistentes no está disponible actualmente para los clústeres de Amazon EMR con varios nodos principales ni para los clústeres de Amazon EMR integrados con Lake Formation. AWS

> **Note**
>
> Para utilizar la autenticación de Kerberos en un clúster, debe configurar un KDC externo. A partir de la versión 5.27.0 de Amazon EMR, puede configurar el cifrado transparente de HDFS en un clúster de Amazon EMR con varios nodos principales. Para más información, consulte Cifrado transparente en el HDFS en Amazon EMR.

Lanzar un clúster de Amazon EMR con varios nodos principales

En este tema se proporcionan detalles de configuración y ejemplos para lanzar un clúster de Amazon EMR con varios nodos principales.

> **Note**
>
> Amazon EMR habilita automáticamente la protección de terminación para todos los clústeres con varios nodos principales y anula cualquier configuración de terminación automática que

se proporcione al crear el clúster. Para cerrar un clúster con varios nodos principales, primero debe modificar los atributos del clúster para deshabilitar la protección contra la terminación. Para obtener instrucciones, consulte Terminar un clúster de Amazon EMR con varios nodos principales.

Requisitos previos

- Puede lanzar un clúster de Amazon EMR con varios nodos principales en subredes públicas y privadas de la VPC. EC2-No se admite la versión clásica. Para lanzar un clúster de Amazon EMR con varios nodos principales en una subred pública, debe habilitar las instancias de esta subred para que reciban una dirección IP pública. Para ello, seleccione Asignación automática IPv4 en la consola o ejecute el siguiente comando. Sustitúyalo por su ID de *22XXXX01* subred.

```
aws ec2 modify-subnet-attribute --subnet-id subnet-22XXXX01 --map-public-ip-on-launch
```

- Para ejecutar Hive, Hue u Oozie en un clúster de Amazon EMR con varios nodos principales, debe crear un metaalmacén externo. Para más información, consulte Configuración de un metaalmacén externo para Hive, Uso de Hue con una base de datos remota en Amazon RDS o Apache Oozie.
- Para utilizar la autenticación de Kerberos en un clúster, debe configurar un KDC externo. Para más información, consulte Configuración de Kerberos en Amazon EMR.

Lanzar un clúster de Amazon EMR con varios nodos principales

Puede lanzar un clúster con varios nodos principales cuando utiliza los grupos de instancias o las flotas de instancias. Al utilizar grupos de instancias con varios nodos principales, debe especificar un valor de recuento de instancias de `3` para el grupo de instancias del nodo principal. Cuando utiliza flotas de instancias con varios nodos principales, debe especificar los valores `TargetOnDemandCapacity` de `3`, `TargetSpotCapacity` de `0` para la flota de instancias principal y `WeightedCapacity` de `1` para cada tipo de instancia que configure para la flota principal.

En los siguientes ejemplos, se muestra cómo lanzar el clúster con la AMI predeterminada o una AMI personalizada tanto con grupos de instancias o con flotas de instancias:

> **Note**
>
> Debe especificar el ID de subred cuando lance un clúster de Amazon EMR con varios nodos principales mediante AWS CLI. Sustituya *22XXXX01* y *22XXXX02* por su ID de subred en los siguientes ejemplos.

Default AMI, instance groups

Example Ejemplo: Lanzamiento de un clúster de grupos de instancias de Amazon EMR con varios nodos principales mediante una AMI predeterminada

```
aws emr create-cluster \
--name "ha-cluster" \
--release-label emr-6.15.0 \
--instance-groups InstanceGroupType=MASTER,InstanceCount=3,InstanceType=m5.xlarge
 InstanceGroupType=CORE,InstanceCount=4,InstanceType=m5.xlarge \
--ec2-attributes
 KeyName=ec2_key_pair_name,InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-22XXXX01
 \
--service-role EMR_DefaultRole \
--applications Name=Hadoop Name=Spark
```

Default AMI, instance fleets

Example Ejemplo: Lanzamiento de un clúster de flotas de instancias de Amazon EMR con varios nodos principales mediante una AMI predeterminada

```
aws emr create-cluster \
--name "ha-cluster" \
--release-label emr-6.15.0 \
--instance-fleets '[
    {
        "InstanceFleetType": "MASTER",
        "TargetOnDemandCapacity": 3,
        "TargetSpotCapacity": 0,
        "LaunchSpecifications": {
            "OnDemandSpecification": {
                "AllocationStrategy": "lowest-price"
            }
        },
        "InstanceTypeConfigs": [
```

```
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.xlarge"
            },
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.2xlarge"
            },
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.4xlarge"
            }
        ],
        "Name": "Master - 1"
    },
    {
        "InstanceFleetType": "CORE",
        "TargetOnDemandCapacity": 5,
        "TargetSpotCapacity": 0,
        "LaunchSpecifications": {
            "OnDemandSpecification": {
                "AllocationStrategy": "lowest-price"
            }
        },
        "InstanceTypeConfigs": [
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.xlarge"
            },
            {
                "WeightedCapacity": 2,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.2xlarge"
            },
            {
                "WeightedCapacity": 4,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.4xlarge"
            }
        ],
```

```
        "Name": "Core - 2"
    }
]' \
 --ec2-attributes '{"InstanceProfile":"EMR_EC2_DefaultRole","SubnetIds":
["subnet-22XXXX01", "subnet-22XXXX02"]}' \
--service-role EMR_DefaultRole \
--applications Name=Hadoop Name=Spark
```

Custom AMI, instance groups

Example Ejemplo: Lanzamiento de un clúster de grupos de instancias de Amazon EMR con varios nodos principales mediante una AMI personalizada

```
aws emr create-cluster \
--name "custom-ami-ha-cluster" \
--release-label emr-6.15.0 \
--instance-groups InstanceGroupType=MASTER,InstanceCount=3,InstanceType=m5.xlarge
 InstanceGroupType=CORE,InstanceCount=4,InstanceType=m5.xlarge \
--ec2-attributes
 KeyName=ec2_key_pair_name,InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-22XXXX01
 \
--service-role EMR_DefaultRole \
--applications Name=Hadoop Name=Spark \
--custom-ami-id ami-MyAmiID
```

Custom AMI, instance fleets

Example Ejemplo: Lanzamiento de un clúster de flotas de instancias de Amazon EMR con varios nodos principales mediante una AMI personalizada

```
aws emr create-cluster \
--name "ha-cluster" \
--release-label emr-6.15.0 \
--instance-fleets '[
    {
        "InstanceFleetType": "MASTER",
        "TargetOnDemandCapacity": 3,
        "TargetSpotCapacity": 0,
        "LaunchSpecifications": {
            "OnDemandSpecification": {
                "AllocationStrategy": "lowest-price"
            }
        },
```

```
        "InstanceTypeConfigs": [
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.xlarge"
            },
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.2xlarge"
            },
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.4xlarge"
            }
        ],
        "Name": "Master - 1"
    },
    {
        "InstanceFleetType": "CORE",
        "TargetOnDemandCapacity": 5,
        "TargetSpotCapacity": 0,
        "LaunchSpecifications": {
            "OnDemandSpecification": {
                "AllocationStrategy": "lowest-price"
            }
        },
        "InstanceTypeConfigs": [
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.xlarge"
            },
            {
                "WeightedCapacity": 2,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.2xlarge"
            },
            {
                "WeightedCapacity": 4,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.4xlarge"
            }
```

```
        ],
        "Name": "Core - 2"
    }
]' \
--ec2-attributes '{"InstanceProfile":"EMR_EC2_DefaultRole","SubnetIds":
["subnet-22XXXX01", "subnet-22XXXX02"]}' \
--service-role EMR_DefaultRole \
--applications Name=Hadoop Name=Spark \
--custom-ami-id ami-MyAmiID
```

Terminar un clúster de Amazon EMR con varios nodos principales

Para terminar un clúster de Amazon EMR con varios nodos principales, debe desactivar la protección de terminación antes de terminar el clúster, tal y como se muestra en el ejemplo siguiente. Sustituya *j-3KVTXXXXXX7UG* por el ID de su clúster.

```
aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-termination-
protected
aws emr terminate-clusters --cluster-id j-3KVTXXXXXX7UG
```

Integración de Amazon EMR con grupos de ubicación EC2

Al lanzar un clúster de múltiples nodos principales de Amazon EMR en Amazon EC2, tiene la opción de utilizar estrategias de grupos de ubicación para especificar cómo quiere que se desplieguen las instancias de nodos principales para protegerlas contra los fallos de hardware.

Las estrategias de grupos de ubicación se admiten a partir de la versión 5.23.0 de Amazon EMR como opción para clústeres con varios nodos principales. Actualmente, la estrategia de grupo de ubicación solo admite los tipos de nodos principales y la estrategia `SPREAD` se aplica a esos nodos. La estrategia `SPREAD` coloca un pequeño grupo de instancias en un hardware subyacente independiente para evitar la pérdida de varios nodos principales en caso de que se produzca un error de hardware. Tenga en cuenta que una solicitud de lanzamiento de instancia podría presentar errores si no hay suficiente hardware único para cumplir con la solicitud. Para obtener más información sobre las estrategias y limitaciones de EC2 ubicación, consulte los grupos de ubicación en la Guía del EC2 usuario de instancias de Linux.

Amazon establece un límite inicial EC2 de 500 clústeres habilitados para la estrategia de grupos de ubicación que se pueden lanzar por región. AWS Ponte en contacto con el servicio de AWS asistencia para solicitar un aumento en el número de grupos de colocación permitidos. Puede

identificar los grupos de EC2 ubicación que Amazon EMR crea mediante el seguimiento del par clave-valor que Amazon EMR asocia a la estrategia de grupos de ubicación de Amazon EMR. Para obtener más información sobre las etiquetas de instancias de EC2 clúster, consulte. Ver instancias de clúster en Amazon EC2

Adjunta la política gestionada por grupos de colocación a Amazon EMRrole

La estrategia de grupos de ubicación requiere una política administrada denominada`AmazonElasticMapReducePlacementGroupPolicy`, que permite a Amazon EMR crear, eliminar y describir grupos de ubicación en Amazon. EC2 Debe asociar `AmazonElasticMapReducePlacementGroupPolicy` al rol de servicio de Amazon EMR antes de lanzar un clúster de Amazon EMR con varios nodos principales.

También puede adjuntar la política `AmazonEMRServicePolicy_v2` gestionada a la función de servicio Amazon EMR en lugar de a la política gestionada por el grupo de colocación. `AmazonEMRServicePolicy_v2`permite el mismo acceso a los grupos de colocación en Amazon EC2 que el`AmazonElasticMapReducePlacementGroupPolicy`. Para obtener más información, consulte Rol de servicio para Amazon EMR (rol de EMR).

La política administrada `AmazonElasticMapReducePlacementGroupPolicy` es el siguiente texto JSON creado y administrado por Amazon EMR.

> **Note**
>
> Como la política `AmazonElasticMapReducePlacementGroupPolicy` gestionada se actualiza automáticamente, es posible que la política que se muestra aquí sí lo esté out-of-date. Utilice la consola AWS de administración para ver la política actual.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Resource": [
        "*"
      ],
```

```
      "Effect": "Allow",
      "Action": [
        "ec2:DeletePlacementGroup",
        "ec2:DescribePlacementGroups"
      ],
      "Sid": "AllowEC2Deleteplacementgroup"
    },
    {
      "Resource": [
        "arn:aws:ec2:*:*:placement-group/pg-*"
      ],
      "Effect": "Allow",
      "Action": [
        "ec2:CreatePlacementGroup"
      ],
      "Sid": "AllowEC2Createplacementgroup"
    }
  ]
}
```

Lanzamiento de un clúster de Amazon EMR con varios nodos principales mediante una estrategia de grupos con ubicación

Para lanzar un clúster de Amazon EMR con varios nodos principales y con una estrategia de grupos con ubicación, asocie la política administrada de grupos con ubicación `AmazonElasticMapReducePlacementGroupPolicy` al rol de Amazon EMR. Para obtener más información, consulte Adjunta la política gestionada por grupos de colocación a Amazon EMRrole.

Cada vez que utiliza este rol para iniciar un clúster de Amazon EMR con varios nodos principales, Amazon EMR intenta lanzar un clúster con la estrategia `SPREAD` aplicada a sus nodos principales. Si utiliza un rol que no tiene la política administrada de grupos con ubicación `AmazonElasticMapReducePlacementGroupPolicy` asociada a él, Amazon EMR intentará lanzar un clúster de Amazon EMR con varios nodos principales sin una estrategia de grupos con ubicación.

Si lanza un clúster de Amazon EMR que tiene varios nodos principales con el `placement-group-configs` parámetro mediante Amazon EMRAPI o CLI, Amazon EMR solo lanza el clúster si EMRrole Amazon tiene adjunta la política de administración de grupos de ubicación. `AmazonElasticMapReducePlacementGroupPolicy` Si Amazon EMRrole no tiene la política adjunta, se produce un error en el inicio del clúster de Amazon EMR con varios nodos principales.

Amazon EMR API

Example Ejemplo: Utilización de una estrategia de grupos con ubicación para lanzar un clúster de grupos de instancias con varios nodos principales desde la API de Amazon EMR

Cuando utilice la RunJobFlow acción para crear un clúster de Amazon EMR con varios nodos principales, defina la `PlacementGroupConfigs` propiedad en lo siguiente. Actualmente, el rol de instancia `MASTER` utiliza automáticamente `SPREAD` como estrategia de grupos de ubicación.

```
{
   "Name":"ha-cluster",
   "PlacementGroupConfigs":[
      {
         "InstanceRole":"MASTER"
      }
   ],
   "ReleaseLabel": emr-6.15.0,
   "Instances":{
      "ec2SubnetId":"subnet-22XXXX01",
      "ec2KeyName":"ec2_key_pair_name",
      "InstanceGroups":[
         {
            "InstanceCount":3,
            "InstanceRole":"MASTER",
            "InstanceType":"m5.xlarge"
         },
         {
            "InstanceCount":4,
            "InstanceRole":"CORE",
            "InstanceType":"m5.xlarge"
         }
      ]
   },
   "JobFlowRole":"EMR_EC2_DefaultRole",
   "ServiceRole":"EMR_DefaultRole"
}
```

- *`ha-cluster`*Sustitúyala por el nombre del clúster de alta disponibilidad.
- *`subnet-22XXXX01`*Sustitúyalo por tu ID de subred.
- Sustituya *`ec2_key_pair_name`* el por el nombre del EC2 key pair de claves de este clúster. EC2 el key pair es opcional y solo es obligatorio si quieres usar SSH para acceder a tu clúster.

AWS CLI

Example Ejemplo: Utilización de una estrategia de grupos con ubicación para lanzar un clúster de flotas de instancias con varios nodos principales desde la AWS Command Line Interface

Cuando utilice la RunJobFlow acción para crear un clúster de Amazon EMR con varios nodos principales, defina la `PlacementGroupConfigs` propiedad en lo siguiente. Actualmente, el rol de instancia `MASTER` utiliza automáticamente `SPREAD` como estrategia de grupos de ubicación.

```
aws emr create-cluster \
--name "ha-cluster" \
--placement-group-configs InstanceRole=MASTER \
--release-label emr-6.15.0 \
--instance-fleets '[
    {
        "InstanceFleetType": "MASTER",
        "TargetOnDemandCapacity": 3,
        "TargetSpotCapacity": 0,
        "LaunchSpecifications": {
            "OnDemandSpecification": {
                "AllocationStrategy": "lowest-price"
            }
        },
        "InstanceTypeConfigs": [
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.xlarge"
            },
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.2xlarge"
            },
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.4xlarge"
            }
        ],
        "Name": "Master - 1"
    },
    {
```

```
        "InstanceFleetType": "CORE",
        "TargetOnDemandCapacity": 5,
        "TargetSpotCapacity": 0,
        "LaunchSpecifications": {
            "OnDemandSpecification": {
                "AllocationStrategy": "lowest-price"
            }
        },
        "InstanceTypeConfigs": [
            {
                "WeightedCapacity": 1,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.xlarge"
            },
            {
                "WeightedCapacity": 2,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.2xlarge"
            },
            {
                "WeightedCapacity": 4,
                "BidPriceAsPercentageOfOnDemandPrice": 100,
                "InstanceType": "m5.4xlarge"
            }
        ],
        "Name": "Core - 2"
    }
]' \
--ec2-attributes '{
    "KeyName": "ec2_key_pair_name",
    "InstanceProfile": "EMR_EC2_DefaultRole",
    "SubnetIds": [
        "subnet-22XXXX01",
        "subnet-22XXXX02"
    ]
}' \
--service-role EMR_DefaultRole \
--applications Name=Hadoop Name=Spark
```

- *ha-cluster*Sustitúyala por el nombre del clúster de alta disponibilidad.
- Sustituya *ec2_key_pair_name* el por el nombre del EC2 key pair de claves de este clúster. EC2 el key pair es opcional y solo es obligatorio si quieres usar SSH para acceder a tu clúster.

- Sustituya *subnet-22XXXX01* y *subnet-22XXXX02* por su IDs subred.

Lanzamiento de un clúster con varios nodos principales sin una estrategia de grupos de ubicación

Para que un clúster con varios nodos principales lance nodos principales sin la estrategia de grupos de ubicación, debe realizar una de las siguientes acciones:

- Eliminar la política gestionada por grupos `AmazonElasticMapReducePlacementGroupPolicy` de colocación de Amazon EMRrole, o
- Lanzar un clúster con varios nodos principales con el parámetro `placement-group-configs` mediante la API o la CLI de Amazon EMR y con `NONE` como estrategia de grupos de ubicación.

Amazon EMR API

Example — Lanzamiento de un clúster con varios nodos principales sin una estrategia de grupos de ubicación mediante la API de Amazon EMR.

Cuando utilice la RunJobFlow acción para crear un clúster con varios nodos principales, defina la `PlacementGroupConfigs` propiedad de la siguiente manera.

```
{
   "Name":"ha-cluster",
   "PlacementGroupConfigs":[
      {
         "InstanceRole":"MASTER",
         "PlacementStrategy":"NONE"
      }
   ],
   "ReleaseLabel":"emr-5.30.1",
   "Instances":{
      "ec2SubnetId":"subnet-22XXXX01",
      "ec2KeyName":"ec2_key_pair_name",
      "InstanceGroups":[
         {
            "InstanceCount":3,
            "InstanceRole":"MASTER",
            "InstanceType":"m5.xlarge"
         },
         {
```

```
            "InstanceCount":4,
            "InstanceRole":"CORE",
            "InstanceType":"m5.xlarge"
         }
      ]
   },
   "JobFlowRole":"EMR_EC2_DefaultRole",
   "ServiceRole":"EMR_DefaultRole"
}
```

- *ha-cluster*Sustitúyala por el nombre del clúster de alta disponibilidad.
- *subnet-22XXXX01*Sustitúyalo por tu ID de subred.
- Sustituya *ec2_key_pair_name* el por el nombre del EC2 key pair de claves de este clúster. EC2 el key pair es opcional y solo es obligatorio si quieres usar SSH para acceder a tu clúster.

Amazon EMR CLI

Example — Lanzamiento de un clúster con varios nodos principales sin una estrategia de grupos de ubicación mediante la CLI de Amazon EMR.

Cuando utilices la RunJobFlow acción para crear un clúster con varios nodos principales, establece la `PlacementGroupConfigs` propiedad de la siguiente manera.

```
aws emr create-cluster \
--name "ha-cluster" \
--placement-group-configs InstanceRole=MASTER,PlacementStrategy=NONE \
--release-label emr-5.30.1 \
--instance-groups InstanceGroupType=MASTER,InstanceCount=3,InstanceType=m5.xlarge
 InstanceGroupType=CORE,InstanceCount=4,InstanceType=m5.xlarge \
--ec2-attributes
 KeyName=ec2_key_pair_name,InstanceProfile=EMR_EC2_DefaultRole,SubnetId=subnet-22XXXX01
 \
--service-role EMR_DefaultRole \
--applications Name=Hadoop Name=Spark
```

- *ha-cluster*Sustitúyala por el nombre del clúster de alta disponibilidad.
- *subnet-22XXXX01*Sustitúyalo por tu ID de subred.
- Sustituya *ec2_key_pair_name* el por el nombre del EC2 key pair de claves de este clúster. EC2 el key pair es opcional y solo es obligatorio si quieres usar SSH para acceder a tu clúster.

Comprobación de la configuración de la estrategia de grupos de ubicación asociada al clúster con varios nodos principales

Puede utilizar la API de descripción del clúster de Amazon EMR para ver la configuración de la estrategia de grupos de ubicación asociada al clúster con varios nodos principales.

Example

```
aws emr describe-cluster --cluster-id "j-xxxxx"
{
   "Cluster":{
      "Id":"j-xxxxx",
      ...
      ...
      "PlacementGroups":[
         {
            "InstanceRole":"MASTER",
            "PlacementStrategy":"SPREAD"
         }
      ]
   }
}
```

Consideraciones y prácticas recomendadas al crear un clúster de Amazon EMR con varios nodos principales

Tenga en cuenta lo siguiente cuando cree un clúster de Amazon EMR con varios nodos principales:

> **Important**
>
> Para lanzar clústeres de EMR de alta disponibilidad con varios nodos principales, le recomendamos encarecidamente que utilice la versión más reciente de Amazon EMR. Esto garantiza que obtenga el nivel más alto de resiliencia y estabilidad para sus clústeres de alta disponibilidad.

- La alta disponibilidad para flotas de instancias es compatible a partir de las versiones 5.36.1, 5.36.2, 6.8.1, 6.9.1, 6.10.1, 6.11.1 y 6.12.0 y posteriores de Amazon EMR. Para grupos de instancias, la alta disponibilidad es compatible a partir de la versión 5.23.0 de Amazon EMR. Para obtener más información, consulte Acerca de las versiones de Amazon EMR.

- En los clústeres de alta disponibilidad, Amazon EMR solo admite el lanzamiento de nodos principales con instancias bajo demanda. Esto garantiza la máxima disponibilidad de su clúster.
- Puede seguir especificando varios tipos de instancias para la flota principal, pero todos los nodos principales de los clústeres de alta disponibilidad se lanzan con el mismo tipo de instancia, incluidas las sustituciones de los nodos principales en mal estado.
- Para continuar funcionando, un clúster de alta disponibilidad con varios nodos principales requiere que dos de cada tres nodos principales estén en buen estado. Como resultado, si dos nodos principales dejan de funcionar de manera simultánea, se producirá un error en el clúster de EMR.
- Todos los clústeres de EMR, incluidos los clústeres de alta disponibilidad, se lanzan en una única zona de disponibilidad. Por lo tanto, no toleran errores de zonas de disponibilidad. En el caso de que deje de funcionar una zona de disponibilidad, se perderá el acceso al clúster.
- Si utiliza un rol o una política de servicio personalizados al lanzar un clúster dentro de una flota de instancias, puede añadir el permiso `ec2:DescribeInstanceTypeOfferings` para que Amazon EMR pueda filtrar las zonas de disponibilidad (AZ) no compatibles. Cuando Amazon EMR filtra los nodos principales AZs que no admiten ningún tipo de instancia, Amazon EMR evita que se produzcan errores en los lanzamientos de clústeres debido a que los tipos de instancias principales no son compatibles. Para obtener más información, consulte Tipos de instancias no admitidos.
- Amazon EMR no garantiza la alta disponibilidad de aplicaciones de código abierto distintas de las especificadas en Aplicaciones admitidas en un clúster de Amazon EMR con varios nodos principales.
- En las versiones desde la 5.23.0 hasta la 5.36.2 de Amazon EMR, solo dos de los tres nodos principales para un clúster de grupos de instancias ejecutan HDFS NameNode.
- En las versiones 6.x y posteriores de Amazon EMR, los tres nodos principales para un grupo de instancias ejecutan HDFS NameNode.

Consideraciones para la configuración la subred:

- Un clúster de Amazon EMR con varios nodos principales solo puede residir en una zona de disponibilidad o subred. Amazon EMR no puede sustituir un nodo principal que ha dejado de funcionar si la subred se utiliza en su totalidad o por encima de su capacidad en caso de que se produzca una conmutación por error. Para evitar esta situación, se recomienda que dedique una subred completa a un clúster de Amazon EMR. Además, debe asegurarse de que haya suficientes direcciones IP privadas disponibles en la subred.

Consideraciones para configurar los nodos principales:

- Para garantizar que los nodos centrales también tengan un alto nivel de disponibilidad, le recomendamos que lance al menos cuatro nodos centrales. Si decide lanzar un clúster más pequeño con tres o menos nodos básicos, configure `dfs.replication parameter` para un mínimo de 2 a fin de que HDFS tenga suficiente replicación DFS. Para más información, consulte Configuración de HDFS.

> **Warning**
>
> 1. Establecer `dfs.replication` en 1 en clústeres con menos de cuatro nodos puede conllevar la pérdida de datos del HDFS si un solo nodo deja de funcionar. Se recomienda que utilice un clúster con al menos cuatro nodos principales para las cargas de trabajo de producción.
> 2. Amazon EMR no permitirá que los clústeres escalen los nodos principales por debajo de `dfs.replication`. Por ejemplo, si `dfs.replication = 2`, el número mínimo de nodos principales es 2.
> 3. Cuando utiliza el escalado administrado, el escalado automático o decide cambiar el tamaño del clúster manualmente, se recomienda que establezca `dfs.replication` en 2 o más.

Consideraciones para configurar alarmas de métricas:

- Amazon EMR no proporciona métricas específicas para las aplicaciones relacionadas con HDFS o YARN. Le recomendamos que configure alarmas para monitorear el recuento de instancias del nodo principal. Configura las alarmas con las siguientes CloudWatch métricas de Amazon: `MultiMasterInstanceGroupNodesRunning``MultiMasterInstanceGroupNodesRunningPercent`
  `oMultiMasterInstanceGroupNodesRequested`. CloudWatch le notificará en caso de que el nodo principal falle o lo reemplace.
  - Si `MultiMasterInstanceGroupNodesRunningPercentage` es inferior al 100% y superior al 50%, es posible que el clúster haya perdido un nodo principal. En esta situación, Amazon EMR intentará sustituir un nodo principal.
  - Si `MultiMasterInstanceGroupNodesRunningPercentage` cae por debajo del 50%, es posible que hayan fallado dos nodos principales. En esta situación, se pierde el cuórum y el clúster no se puede recuperar. Debe migrar manualmente los datos de este clúster.

Para más información, consulte Configuración de alarmas en métricas.

EMR se agrupa en AWS Outposts

A partir de Amazon EMR 5.28.0, puede crear y ejecutar clústeres de EMR en ellos. AWS Outposts AWS Outposts habilita AWS los servicios, la infraestructura y los modelos operativos nativos en instalaciones locales. En AWS Outposts los entornos, puede usar las mismas AWS APIs herramientas e infraestructura que usa en la AWS nube. Amazon EMR on AWS Outposts es ideal para cargas de trabajo de baja latencia que deben ejecutarse muy cerca de datos y aplicaciones locales. Para obtener más información al respecto AWS Outposts, consulte AWS Outposts la Guía del usuario.

Requisitos previos

Estos son los requisitos previos para utilizar Amazon EMR en AWS Outposts:

- Debe haberlo instalado y configurado AWS Outposts en su centro de datos local.
- Debe disponer de una conexión de red fiable entre su entorno de Outpost y una AWS región.
- Debe tener capacidad suficiente para los tipos de instancias compatibles de Amazon EMR disponibles en su Outpost.

Limitaciones

A continuación, se indican las limitaciones de uso de Amazon EMR en AWS Outposts:

- Las instancias bajo demanda son la única opción compatible con las EC2 instancias de Amazon. Las instancias de spot no están disponibles para Amazon EMR en AWS Outposts.
- Si necesita volúmenes de almacenamiento adicionales de Amazon EBS, solo se admiten SSD de uso general (GP2).
- Si lo usa AWS Outposts con las versiones 5.28 a 6.x de Amazon EMR, solo puede usar buckets de S3 que almacenen objetos en un contenedor que usted especifique. Región de AWS Con Amazon EMR 7.0.0 y versiones posteriores, Amazon EMR on también AWS Outposts es compatible con el prefijo del cliente del S3A sistema de archivos. `s3a://`
- Únicamente los siguientes tipos de instancias son admitidas por Amazon EMR en AWS Outposts:

Clase de instancia	Tipos de instancias
Uso general	`m5.xlarge` | `m5.2xlarge` | `m5.4xlarge` | `m5.12xlarge` | `m5.24xlarge` | `m5d.xlarge` | `m5d.2xlarge` | `m5d.4xlarge` | `m5d.12xlarge` | `m5d.24xlarge`
Optimizada para computación	`c5.xlarge` | `c5.2xlarge` | `c5.4xlarge` | `c5.18xlarge` | `c5d.xlarge` | `c5d.2xlarge` | `c5d.4xlarge` | `c5d.18xlarge`
Optimizada para memoria	`r5.xlarge` | `r5.2xlarge` | `r5.4xlarge` | `r5.12xlarge` | `r5d.xlarge` | `r5d.2xlarge` | `r5d.4xlarge` | `r5d.12xlarge` | `r5d.24xlarge`
Con optimización del almacenamiento	`i3en.xlarge` | `i3en.2xlarge` | `i3en.3xlarge` | `i3en.6xlarge` | `i3en.12xlarge` | `i3en.24xlarge`

Consideraciones sobre la conectividad de red

- Si se pierde la conectividad de red entre su Outpost y su AWS región, sus clústeres seguirán funcionando. Sin embargo, no podrá crear nuevos clústeres ni realizar nuevas acciones en clústeres existentes hasta que se restablezca la conectividad. En caso de errores en la instancia, la instancia no se reemplazará automáticamente. Además, se retrasarán acciones como añadir pasos a un clúster en ejecución, comprobar el estado de ejecución de los pasos y enviar CloudWatch métricas y eventos.
- Te recomendamos que proporciones una conectividad de red fiable y de alta disponibilidad entre tu puesto de avanzada y la AWS región. Si se pierde la conectividad de red entre tu Outpost y su AWS región durante más de unas horas, los clústeres que tengan habilitada la protección de terminales seguirán funcionando y los clústeres que la hayan desactivado pueden terminar.
- Si la conectividad de red se ve afectada debido a un mantenimiento rutinario, se recomienda habilitar proactivamente la protección de terminación. De manera más general, la interrupción de la conectividad significa que no se podrá acceder a ninguna dependencia externa que no esté accesible localmente en la instancia de Outpost o en la red del cliente. Esto incluye Amazon S3,

DynamoDB (utilizado con la vista de coherencia de EMRFS) y Amazon RDS, si se utiliza una instancia regional para un clúster de Amazon EMR con varios nodos principales.

Crear un clúster de Amazon EMR en AWS Outposts

Crear un clúster de Amazon EMR en AWS Outposts es similar a crear un clúster de Amazon EMR en la nube. AWS Al crear un clúster de Amazon EMR en AWS Outposts, debe especificar una EC2 subred de Amazon asociada a su Outpost.

Una Amazon VPC puede abarcar todas las zonas de disponibilidad de una AWS región. AWS Outposts son extensiones de las zonas de disponibilidad y puede ampliar una Amazon VPC en una cuenta para abarcar varias zonas de disponibilidad y las ubicaciones de Outpost asociadas. Al configurar su Outpost, le asocia un grupo de subredes para ampliar su entorno de VPC regional a sus instalaciones. Las instancias de Outpost y los servicios relacionados aparecen como parte de su VPC regional, de manera similar a una zona de disponibilidad con subredes asociadas. Para obtener más información, consulte la Guía del usuario de AWS Outposts.

Consola

Para crear un nuevo clúster de Amazon EMR AWS Outposts con AWS Management Console, especifique una EC2 subred de Amazon que esté asociada a su Outpost.

Console

Para crear un clúster en la consola AWS Outposts

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Configuración del clúster, seleccione Grupos de instancias o Flotas de instancias. A continuación, elija un tipo de instancia en el menú desplegable Elegir tipo de EC2 instancia o seleccione Acciones y elija Agregar volúmenes de EBS. Amazon EMR on AWS Outposts admite tipos limitados de volúmenes e instancias de Amazon EBS.
4. En Redes, seleccione una EC2 subred con un ID de Outpost en este formato: op-123456789.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

CLI

Para crear un clúster con AWS OutpostsAWS CLI

- Para crear un nuevo clúster de Amazon EMR AWS Outposts con AWS CLI, especifique una EC2 subred que esté asociada a su Outpost, como en el siguiente ejemplo. *subnet-22XXXX01*Sustitúyalo por tu propio ID de EC2 subred de Amazon.

```
aws emr create-cluster \
--name "Outpost cluster" \
--release-label emr-7.10.0 \
--applications Name=Spark \
--ec2-attributes KeyName=myKey SubnetId=subnet-22XXXX01 \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Clústeres de EMR en Zonas Locales AWS

A partir de la versión 5.28.0 de Amazon EMR, puede crear y ejecutar clústeres de Amazon EMR en una subred de Zonas AWS Locales como extensión lógica de una región que admite Zonas Locales. AWS Una zona local permite que las funciones de Amazon EMR y un subconjunto de AWS servicios, como los servicios de procesamiento y almacenamiento, se ubiquen más cerca de los usuarios para proporcionar un acceso de latencia muy baja a las aplicaciones que se ejecutan localmente. Para obtener una lista de las zonas locales disponibles, consulte Zonas locales de AWS. Para obtener información sobre cómo acceder a las zonas AWS locales disponibles, consulte Regiones, zonas de disponibilidad y zonas locales.

Tipos de instancias admitidas

Los siguientes tipos de instancias están disponibles para los clústeres de Amazon EMR en zonas locales. La disponibilidad de los distintos tipos de instancias puede variar según la región.

Clase de instancia	Tipos de instancias
Uso general	m5.xlarge | m5.2xlarge | m5.4xlarge | m5.12xlarge | m5.24xlarge | m5d.xlarge | m5d.2xlarge | m5d.4xlarge | m5d.12xlarge | m5d.24xlarge

Clase de instancia	Tipos de instancias
Optimizada para computación	c5.xlarge | c5.2xlarge | c5.4xlarge | c5.9xlarge | c5.18xlarge | c5d.xlarge | c5d.2xlarge | c5d.4xlarge| c5d.9xlarge | c5d.18xlarge
Optimizada para memoria	r5.xlarge | r5.2xlarge | r5.4xlarge | r5.12xlarge | r5d.xlarge | r5d.2xlarge | r5d.4xlarge | r5d.12xlarge | r5d.24xlarge
Con optimización del almacenamiento	i3en.xlarge | i3en.2xlarge | i3en.3xlarge | i3en.6xlarge | i3en.12xlarge | i3en.24xlarge

Creación de un clúster de Amazon EMR en zonas locales

Cree un clúster de Amazon EMR en las zonas AWS locales lanzando el clúster de Amazon EMR en una subred de Amazon VPC asociada a una zona local. Puede acceder al clúster utilizando el nombre de la zona local; por ejemplo, us-west-2-lax-1a en la consola de Oeste de EE. UU. (Oregón).

Actualmente, las Zonas Locales no admiten los Amazon EMR Notebooks ni las conexiones directas a Amazon EMR mediante el punto de enlace de la interfaz VPC ().AWS PrivateLink

Console

Para crear un clúster en una zona local con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Redes, seleccione una EC2 subred con un identificador de zona local en este formato: subnet 123abc | us-west-2-lax-1a.
4. Elija un tipo de instancia o agregue volúmenes de almacenamiento de Amazon EBS para grupos de instancias uniformes o flotas de instancias.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

CLI

Para crear un clúster en una zona local con AWS CLI

- Utilice el comando create-cluster junto con el comando SubnetId para la zona local, como se muestra en el siguiente ejemplo. Sustituya la subnet-22 por XXXX1234567 la zona local SubnetId y sustituya las demás opciones según sea necesario. Para obtener más información, consulte https://docs.aws.amazon.com/cli/latest/reference/emr/create-cluster.html.

```
aws emr create-cluster \
--name "Local Zones cluster" \
--release-label emr-5.29.0 \
--applications Name=Spark \
--ec2-attributes KeyName=myKey,SubnetId=subnet-22XXXX1234567 \
--instance-type m5.xlarge --instance-count 3 --use-default-roles
```

Configuración de Docker para su uso con clústeres de Amazon EMR

Amazon EMR 6.x es compatible con Hadoop 3, lo que permite NodeManager a YARN lanzar contenedores directamente en el clúster de Amazon EMR o dentro de un contenedor Docker. Los contenedores de Docker proporcionan entornos de ejecución personalizados en los que se ejecuta el código de aplicación. El entorno de ejecución personalizado está aislado del entorno de ejecución del YARN y de otras aplicaciones. NodeManager

Los contenedores de Docker pueden incluir bibliotecas especiales utilizadas por la aplicación y proporcionar diferentes versiones de las herramientas y las bibliotecas nativas, como R y Python. Puede utilizar las herramientas de Docker conocidas para definir dependencias en tiempo de ejecución y bibliotecas para sus aplicaciones.

Los clústeres de Amazon EMR 6.x se han configurado de forma predeterminada para permitir que las aplicaciones de YARN, como Spark, se ejecuten utilizando contenedores de Docker. Para personalizar la configuración del contenedor, edite las opciones de soporte de Docker definidas en los archivos `container-executor.cfg` y `yarn-site.xml` disponibles en el directorio `/etc/hadoop/conf`. Para obtener información detallada sobre cada opción de configuración y cómo se utiliza, consulte Lanzamiento de aplicaciones mediante contenedores de Docker.

Si lo desea, puede usar Docker cuando envíe un trabajo. Use las siguientes variables para especificar el tiempo de ejecución de Docker y la imagen de Docker.

- `YARN_CONTAINER_RUNTIME_TYPE=docker`
- `YARN_CONTAINER_RUNTIME_DOCKER_IMAGE={DOCKER_IMAGE_NAME}`

Cuando usa contenedores de Docker para ejecutar sus aplicaciones de YARN, YARN descarga la imagen de Docker que se ha especificado al enviar el trabajo. Para que YARN resuelva esta imagen de Docker, debe configurarse con un registro de Docker. Las opciones de configuración de un registro de Docker dependen de si el clúster se ha implementado mediante una subred pública o privada.

Registros de Docker

Un registro de Docker es un sistema de almacenamiento y distribución para imágenes de Docker. En el caso de Amazon EMR, se recomienda usar Amazon ECR, que es un registro de contenedores de Docker completamente administrado que le permite crear sus propias imágenes personalizadas y alojarlas en una arquitectura escalable y de alta disponibilidad.

Consideraciones sobre la implementación

Los registros de Docker requieren acceso a la red desde cada host del clúster. Esto se debe a que cada host descarga imágenes del registro de Docker cuando la aplicación de YARN se ejecuta en el clúster. Estos requisitos de conectividad de red pueden limitar la elección del registro de Docker, dependiendo de si el clúster de Amazon EMR se ha implementado en una subred pública o privada.

Public subnet (Subred pública)

Cuando los clústeres EMR se implementan en una subred pública, los nodos que ejecutan YARN NodeManager pueden acceder directamente a cualquier registro disponible en Internet.

Subred privada

Cuando los clústeres de EMR se implementan en una subred privada, los nodos que ejecutan YARN NodeManager no tienen acceso directo a Internet. Las imágenes de Docker se pueden alojar en Amazon ECR y acceder a ellas a través de. AWS PrivateLink

Para obtener más información sobre cómo permitir el acceso AWS PrivateLink a Amazon ECR en un escenario de subred privada, consulte Configuración de AWS PrivateLink Amazon ECS y Amazon ECR.

Configuración de registros de Docker

Para utilizar registros de Docker con Amazon EMR, debe configurar Docker de tal forma que confíe en el registro específico que desea utilizar para resolver las imágenes de Docker. Los registros de confianza predeterminados son locales (privados) y de Centos. Para utilizar otros repositorios públicos o Amazon ECR, puede anular la configuración de `docker.trusted.registries` en `/etc/hadoop/conf/container-executor.cfg` mediante la API de clasificación de EMR con la clave de clasificación `container-executor`.

En el ejemplo siguiente se muestra cómo configurar el clúster de tal forma que confíe tanto en un repositorio público denominado `your-public-repo`, como en un punto de enlace de registro de ECR, `123456789123.dkr.ecr.us-east-1.amazonaws.com`. Si utiliza ECR, sustituya este punto de enlace por su punto de enlace de ECR específico.

```
[
  {
    "Classification": "container-executor",
    "Configurations": [
        {
            "Classification": "docker",
            "Properties": {
                "docker.trusted.registries": "local,centos,your-public-
repo,123456789123.dkr.ecr.us-east-1.amazonaws.com",
                "docker.privileged-containers.registries": "local,centos,your-public-
repo,123456789123.dkr.ecr.us-east-1.amazonaws.com"
            }
        }
    ]
  }
]
```

Para lanzar un clúster de Amazon EMR 6.0.0 con esta configuración mediante AWS Command Line Interface (AWS CLI), cree un archivo con el nombre del contenido de la configuración `container-executor.json` JSON del contenedor-ejecutor anterior. A continuación, utilice los siguientes comandos para lanzar el clúster.

```
export KEYPAIR=<Name of your Amazon EC2 key-pair>
export SUBNET_ID=<ID of the subnet to which to deploy the cluster>
export INSTANCE_TYPE=<Name of the instance type to use>
export REGION=<Region to which to deploy the cluster>
```

```
aws emr create-cluster \
    --name "EMR-6.0.0" \
    --region $REGION \
    --release-label emr-6.0.0 \
    --applications Name=Hadoop Name=Spark \
    --service-role EMR_DefaultRole \
    --ec2-attributes KeyName=$KEYPAIR,InstanceProfile=EMR_EC2_DefaultRole,SubnetId=
$SUBNET_ID \
    --instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=
$INSTANCE_TYPE InstanceGroupType=CORE,InstanceCount=2,InstanceType=$INSTANCE_TYPE \
    --configuration file://container-executor.json
```

Configuración de YARN para acceder a Amazon ECR en EMR 6.0.0 y versiones anteriores

Si es la primera vez que usa Amazon ECR, siga las instrucciones de Introducción a Amazon ECR y compruebe que dispone de acceso a Amazon ECR desde cada instancia del clúster de Amazon EMR.

En EMR 6.0.0 y versiones anteriores, para acceder a Amazon ECR mediante el comando de Docker, antes debe generar unas credenciales. Para verificar que YARN pueda acceder a las imágenes desde Amazon ECR, utilice la variable de entorno contenedora `YARN_CONTAINER_RUNTIME_DOCKER_CLIENT_CONFIG` para pasar una referencia a las credenciales que ha generado.

Ejecute el siguiente comando en uno de los nodos principales para obtener la línea de inicio de sesión de su cuenta de ECR.

```
aws ecr get-login --region us-east-1 --no-include-email
```

El comando `get-login` genera el comando correcto de la CLI de Docker que hay que ejecutar para crear las credenciales. Copie y ejecute la salida desde `get-login`.

```
sudo docker login -u AWS -p <password> https://<account-id>.dkr.ecr.us-
east-1.amazonaws.com
```

Este comando genera un archivo `config.json` en la carpeta `/root/.docker`. Copie este archivo en HDFS para que los trabajos enviados al clúster puedan usarlo para autenticarse en Amazon ECR.

Ejecute los siguientes comandos para copiar el archivo `config.json` en su directorio de inicio.

```
mkdir -p ~/.docker
sudo cp /root/.docker/config.json ~/.docker/config.json
sudo chmod 644 ~/.docker/config.json
```

Ejecute los siguientes comandos para poner el archivo config.json en HDFS, de tal forma que puedan usarlo los trabajos que se ejecutan en el clúster.

```
hadoop fs -put ~/.docker/config.json /user/hadoop/
```

YARN puede acceder a ECR como un registro de imágenes de Docker y extraer contenedores durante la ejecución del trabajo.

Después de configurar los registros de Docker y YARN, puede ejecutar las aplicaciones de YARN utilizando contenedores de Docker. Para más información, consulte Ejecutar aplicaciones de Spark con Docker mediante Amazon EMR 6.0.0.

En EMR 6.1.0 y versiones posteriores, no es necesario configurar manualmente la autenticación en Amazon ECR. Si se detecta un registro de Amazon ECR en la clave de clasificación `container-executor`, se activa la característica de autenticación automática de Amazon ECR y YARN se encarga del proceso de autenticación cuando se envía un trabajo de Spark con una imagen de ECR. Para confirmar si la autenticación automática está habilitada, compruebe `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled` en yarn-site. La autenticación automática está habilitada y la configuración de autenticación de YARN se establece en `true` si `docker.trusted.registries` contiene una URL de registro de ECR.

Requisitos previos para utilizar la autenticación automática en Amazon ECR

- EMR versión 6.1.0 o posterior
- El registro de ECR incluido en la configuración se encuentra en la misma región que el clúster
- Rol de IAM con permisos para obtener el token de autorización y extraer cualquier imagen

Consulte Configuración con Amazon ECR para más información.

Cómo habilitar la autenticación automática

Siga Configuración de registros de Docker para establecer un registro de Amazon ECR como registro de confianza y asegúrese de que el repositorio de Amazon ECR y el clúster estén en la misma región.

Para habilitar esta característica incluso cuando el registro de ECR no esté establecido en el registro de confianza, utilice la clasificación de configuración para establecer `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled` en `true`.

Cómo desactivar la autenticación automática

De forma predeterminada, la autenticación automática se deshabilita si no se detecta ningún registro de Amazon ECR en el registro de confianza.

Para desactivar la autenticación automática, incluso cuando el registro de Amazon ECR se haya establecido en el registro de confianza, utilice la clasificación de configuración para establecer `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled` en `false`.

Cómo comprobar si la autenticación automática está habilitada en un clúster

En el nodo maestro, utilice un editor de texto como `vi` para ver el contenido del archivo: `vi /etc/hadoop/conf.empty/yarn-site.xml`. Compruebe el valor de `yarn.nodemanager.runtime.linux.docker.ecr-auto-authentication.enabled`.

Control de la finalización de clústeres de Amazon EMR

En esta sección se describen las opciones para cerrar los clústeres de Amazon EMR. Abarca la terminación automática y la protección contra la terminación, y la forma en que interactúan con otras características de Amazon EMR.

Puede cerrar un clúster de Amazon EMR de las siguientes maneras:

- Terminación tras la ejecución del último paso: cree un clúster transitorio que se cierre una vez completados todos los pasos.
- Terminación automática (después de la inactividad): cree un clúster con una política de terminación automática que se cierre después de un tiempo de inactividad específico. Para obtener más información, consulte Uso de una política de finalización automática para la limpieza de clústeres de Amazon EMR..
- Terminación manual: cree un clúster de ejecución prolongada que continúe ejecutándose hasta que lo termine deliberadamente. Para obtener información sobre cómo terminar un clúster de forma manual, consulte Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera..

También puede configurar la protección de terminación en un clúster para evitar el cierre de las EC2 instancias por accidente o error.

Cuando Amazon EMR cierra el clúster, se cierran todas las EC2 instancias de Amazon del clúster. Los datos del almacén de instancias y de los volúmenes de EBS ya no están disponibles ni se pueden recuperar. La comprensión y administración de la terminación de los clústeres es fundamental para desarrollar una estrategia que permita administrar y preservar los datos escribiéndolos en Amazon S3 y equilibrando los costos.

Temas

- Configuración de un clúster de Amazon EMR para que continúe o finalice después de la ejecución de pasos
- Uso de una política de finalización automática para la limpieza de clústeres de Amazon EMR.
- Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental

Configuración de un clúster de Amazon EMR para que continúe o finalice después de la ejecución de pasos

En este tema se explican las diferencias entre utilizar un clúster de ejecución prolongada y crear un clúster transitorio que se cierre después de ejecutar el último paso. También se explica cómo configurar la ejecución de pasos de un clúster.

Crear un clúster en ejecución prolongada

De forma predeterminada, los clústeres que se crean con la consola o AWS CLI son de larga duración. Los clústeres de ejecución prolongada siguen funcionando, aceptando trabajo y acumulando gastos hasta que tome medidas para cerrarlos.

Un clúster de ejecución prolongada es eficaz en las siguientes situaciones:

- Cuando necesite consultar datos de forma interactiva o automática.
- Cuando necesite interactuar con aplicaciones de macrodatos alojadas en el clúster de forma continua.
- Cuando procesa periódicamente un conjunto de datos tan grande o con tanta frecuencia que resulta poco eficiente lanzar clústeres nuevos y cargar los datos cada vez.

También puede configurar la protección de terminación en un clúster de ejecución prolongada para evitar el cierre de EC2 las instancias por accidente o error. Para obtener más información, consulte Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental.

> **Note**
>
> Amazon EMR habilita automáticamente la protección contra la terminación para todos los clústeres con varios nodos principales y anula cualquier configuración de ejecución de pasos que proporcione al crear el clúster. Puede deshabilitar la protección contra la terminación después de que se haya lanzado el clúster. Consulte Configuración de la protección de terminación para ejecutar clústeres. Para cerrar un clúster con varios nodos principales, primero debe modificar los atributos del clúster para deshabilitar la protección contra la terminación. Para obtener instrucciones, consulte Terminar un clúster de Amazon EMR con varios nodos principales.

Configuración de un clúster para que termine tras la ejecución de pasos

Al configurar la terminación tras la ejecución de pasos, el clúster se inicia, ejecuta las acciones de arranque y, a continuación, ejecuta los pasos que especifique. En cuanto se complete el último paso, Amazon EMR finaliza las instancias de Amazon del clúster. EC2 Los clústeres que lance con la API de Amazon EMR tienen habilitada la ejecución de pasos de forma predeterminada.

La terminación posterior a la ejecución de pasos es eficaz para los clústeres que realizan una tarea de procesamiento periódica, como un procesamiento de datos diario. La ejecución de pasos también lo ayuda a garantizar que solo se le facture el tiempo necesario para procesar sus datos. Para obtener más información acerca de los pasos, consulte Envío del trabajo a un clúster de Amazon EMR.

Console

Para activar la finalización tras la ejecución de pasos con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.

3. En Pasos, seleccione Agregar paso. En el cuadro de diálogo Agregar paso, introduzca los valores en los campos adecuados. Las opciones varían según el tipo de paso. Para agregar el paso y salir del cuadro de diálogo, elija Agregar paso.
4. En Terminación del clúster, seleccione la casilla Terminar el clúster tras completar el último paso.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para activar la terminación tras la ejecución de los pasos con el AWS CLI

- Especifique el parámetro `--auto-terminate` cuando utilice el comando `create-cluster` para crear un clúster transitorio.

  El siguiente ejemplo muestra el uso del parámetro `--auto-terminate`. Puede escribir el siguiente comando y *`myKey`* sustituirlo por el nombre del par de EC2 claves.

  > **Note**
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.10.0 \
--applications Name=Hive Name=Pig --use-default-roles --ec2-attributes
 KeyName=myKey \
--steps Type=PIG,Name="Pig Program",ActionOnFailure=CONTINUE,\
Args=[-f,s3://amzn-s3-demo-bucket/scripts/pigscript.pig,-p,\
INPUT=s3://amzn-s3-demo-bucket/inputdata/,-p,OUTPUT=s3://amzn-s3-demo-bucket/
outputdata/,\
$INPUT=s3://amzn-s3-demo-bucket/inputdata/,$OUTPUT=s3://amzn-s3-demo-bucket/
outputdata/]
--instance-type m5.xlarge --instance-count 3 --auto-terminate
```

API

Para desactivar la finalización tras la ejecución de pasos con la API de Amazon EMR en el lanzamiento del clúster

1. Cuando utilice la RunJobFlowacción para crear un clúster, defina la KeepJobFlowAliveWhenNoStepspropiedad `enfalse`.
2. Para cambiar su configuración de finalización tras la ejecución de pasos con la API de Amazon EMR tras el lanzamiento del clúster:

   Usa SetKeepJobFlowAliveWhenNoSteps la acción.

Uso de una política de finalización automática para la limpieza de clústeres de Amazon EMR.

Una política de terminación automática le permite orquestar la limpieza de los clústeres sin necesidad de supervisar ni terminar manualmente los clústeres no utilizados. Cuando agrega una política de terminación automática a un clúster, debe especificar la cantidad de tiempo de inactividad tras el cual el clúster debe cerrarse automáticamente.

Según la versión de lanzamiento, Amazon EMR utiliza diferentes criterios para marcar un clúster como inactivo. En la siguiente tabla se describe cómo Amazon EMR determina la inactividad del clúster.

Cuando utiliza…	Un clúster se considera inactivo cuando…
Amazon EMR, versiones 5.34.0 y posteriores, y 6.4.0 y posteriores	• No hay aplicaciones de YARN activas • La utilización del HDFS es inferior al 10 % • No hay conexiones activas a cuadernos de EMR o a EMR Studio • No se utilizan interfaces de usuario de aplicaciones en el clúster • No hay pasos pendientes

Cuando utiliza…	Un clúster se considera inactivo cuando…
Amazon EMR, versiones 5.30.0 a 5.33.0 y 6.1.0 a 6.3.0	• No hay aplicaciones de YARN activas • El clúster no tiene trabajos de Spark activos **Note** Amazon EMR marca un clúster como inactivo y puede terminarlo automátic amente incluso si tiene un kernel de Python3 activo. Esto se debe a que al ejecutar un kernel de Python3 no se envía ningún trabajo de Spark al clúster. Para utilizar la terminaci ón automática con un kernel de Python3, le recomendamos que utilice la versión 6.4.0 o posterior de Amazon EMR.

Note

Las versiones 6.4.0 y posteriores de Amazon EMR admiten un archivo en el clúster para detectar la actividad en el nodo principal: `/emr/metricscollector/isbusy`. Cuando utiliza un clúster para ejecutar scripts de intérprete de comandos o aplicaciones que no son de YARN, puede tocar o actualizar periódicamente `isbusy` para indicar a Amazon EMR que el clúster no está inactivo.

Puede asociar una política de terminación automática al crear un clúster o agregar una política a un clúster existente. Para cambiar o deshabilitar la terminación automática, puede actualizar o eliminar la política.

Consideraciones

Tenga en cuenta las siguientes características y limitaciones antes de utilizar una política de terminación automática:

- A continuación Regiones de AWS, la terminación automática de Amazon EMR está disponible con Amazon EMR 6.14.0 y versiones posteriores:
  - Europa (España) (eu-south-2)
- A continuación Regiones de AWS, la terminación automática de Amazon EMR está disponible con Amazon EMR 5.30.0 y 6.1.0 y versiones posteriores:
  - Este de EE. UU. (Norte de Virginia) (us-east-1)
  - Este de EE. UU. (Ohio) (us-east-2)
  - Oeste de EE. UU. (Oregón) (us-west-2)
  - EE. UU. Oeste (Norte de California) (us-west-1)
  - África (Ciudad del Cabo) (af-south-1)
  - Asia-Pacífico (Hong Kong) (ap-east-1)
  - Asia Pacífico (Bombay) (ap-south-1)
  - Asia Pacífico (Hyderabad) (ap-south-2)
  - Asia-Pacífico (Seúl) (ap-northeast-2)
  - Asia Pacific (Osaka) (ap-northeast-3)
  - Asia-Pacífico (Singapur) (ap-southeast-1)
  - Asia-Pacífico (Sídney) (ap-southeast-2)
  - Asia-Pacífico (Yakarta) (ap-southeast-3)
  - Asia-Pacífico (Tokio) (ap-northeast-1)
  - Canadá (centro) (ca-central-1)
  - América del Sur (São Paulo) (sa-east-1)
  - Europa (Fráncfort) (eu-central-1)
  - Europa (Zúrich) (eu-central-2)
  - Europa (Irlanda) (eu-west-1)
  - Europa (Londres) (eu-west-2)
  - UE (Milán) (eu-south-1)

  - UE (París) (eu-west-3)

  - Europa (Estocolmo) (eu-north-1)
  - Israel (Tel Aviv) (il-central-1)
  - Medio Oriente (EAU) (me-central-1)
  - China (Pekín) (cn-north-1)
  - China (Ningxia) (cn-northwest-1)
  - AWS GovCloud (EE. UU.-Este) (-1) us-gov-east
  - AWS GovCloud (EEUU-Oeste) (us-gov-west-1)
- El tiempo de espera de inactividad se establece de forma predeterminada en 60 minutos (una hora) si no se especifica una cantidad. Puede especificar un tiempo de espera de inactividad mínimo de un minuto y un tiempo de espera de inactividad máximo de 7 días.
- Con las versiones 6.4.0 y posteriores de Amazon EMR, la terminación automática se habilita de forma predeterminada al crear un clúster nuevo con la consola de Amazon EMR.
- Amazon EMR publica Amazon CloudWatch métricas de alta resolución cuando habilita la terminación automática de un clúster. Puede usar estas métricas para realizar un seguimiento de la actividad y la inactividad del clúster. Para obtener más información, consulte Métricas de capacidad del clúster.
- La terminación automática no se admite cuando se utilizan aplicaciones que no estén basadas en Yarn, como Presto, Trino o. HBase
- Para utilizar la terminación automática, el proceso de recopilación de métricas debe poder conectarse al punto de conexión de la API pública para la terminación automática en API Gateway. Si utilizas un nombre DNS privado con Amazon Virtual Private Cloud, la terminación automática no funcionará correctamente. Para garantizar que la terminación automática funcione, se recomienda que realice una de las siguientes acciones:
  - Elimine el punto de conexión de VPC de la interfaz de API Gateway de su Amazon VPC.
  - Siga las instrucciones de ¿Por qué aparece el error HTTP 403 Forbidden al conectarme a mi API Gateway APIs desde una VPC? para deshabilitar la configuración del nombre DNS privado.
  - En su lugar, lance su clúster en una subred privada. Para más información, consulte el tema sobre Subredes privadas.
- (EMR 5.30.0 y versiones posteriores) Si quita la regla de salida predeterminada Permitir todo a 0.0.0.0/ para el grupo de seguridad principal, debe agregar una regla que permita la conectividad TCP de salida a su grupo de seguridad para el acceso al servicio en el puerto 9443. El grupo de seguridad para el acceso al servicio también debe permitir el tráfico TCP entrante en el puerto 9443 desde el grupo de seguridad principal. Para más información sobre la configuración

de grupos de seguridad, consulte Grupo de seguridad administrado por Amazon EMR para la instancia principal (subredes privadas).

Permisos para utilizar la terminación automática

Antes de poder aplicar y administrar políticas de terminación automática para Amazon EMR, debe asociar los permisos que se enumeran en el siguiente ejemplo de política de permisos de IAM a los recursos de IAM que administran su clúster de EMR.

```
{
  "Version": "2012-10-17",
  "Statement": {
      "Sid": "AllowAutoTerminationPolicyActions",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:PutAutoTerminationPolicy",
        "elasticmapreduce:GetAutoTerminationPolicy",
        "elasticmapreduce:RemoveAutoTerminationPolicy"
      ],
      "Resource": "<your-resources>"
    }
}
```

Asociación, actualización o eliminación de una política de terminación automática

En esta sección se incluyen instrucciones que le ayudarán a asociar, actualizar o eliminar una política de terminación automática de un clúster de Amazon EMR. Antes de trabajar con políticas de terminación automática, asegúrese de tener los permisos de IAM necesarios. Consulte Permisos para utilizar la terminación automática.

Console

Para asociar una política de finalización automática al crear un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Terminación del clúster, seleccione Terminar el clúster después del tiempo de inactividad.

4. Especifique el número de horas y minutos de inactividad que pueden transcurrir antes de que el clúster termine automáticamente. El tiempo de espera de inactividad predeterminado es de 1 hora.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

Para asociar, actualizar o eliminar una política de finalización automática en un clúster en ejecución con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. En la pestaña Propiedades de la página de detalles del clúster, busque Terminación del clúster y seleccione Editar.
4. Seleccione o anule la selección de Habilitar la terminación automática para activar o desactivar la característica. Si activa la terminación automática, especifique el número de horas y minutos de inactividad que pueden transcurrir antes de que el clúster termine automáticamente. A continuación, seleccione Guardar cambios para confirmar.

AWS CLI

Antes de comenzar

Antes de trabajar con políticas de terminación automática, se recomienda que actualice a la versión más reciente de la AWS CLI. Para obtener instrucciones, consulte Instalación, actualización y desinstalación de la AWS CLI.

Para asociar o actualizar una política de terminación automática mediante AWS CLI

- Puede utilizar el comando `aws emr put-auto-termination-policy` para asociar o actualizar una política de terminación automática en un clúster.

  En el siguiente ejemplo se especifican 3600 segundos para. *`IdleTimeout`* Si no lo especifica*`IdleTimeout`*, el valor predeterminado es una hora.

```
aws emr put-auto-termination-policy \
```

```
--cluster-id <your-cluster-id> \
--auto-termination-policy IdleTimeout=3600
```

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

También puede especificar un valor para `--auto-termination-policy` cuando utilice el comando `aws emr create-cluster`. Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte la Referencia de AWS CLI comandos.

Para eliminar una política de rescisión automática con la AWS CLI

- Utilice el comando `aws emr remove-auto-termination-policy` para eliminar una política de terminación automática de un clúster. Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte la Referencia de AWS CLI comandos.

  ```
  aws emr remove-auto-termination-policy --cluster-id <your-cluster-id>
  ```

Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental

La protección de finalización protege sus clústeres de la finalización accidental, lo que puede resultar especialmente útil para clústeres de larga ejecución que procesan cargas de trabajo críticas. Si la protección de terminación está habilitada en un clúster de ejecución prolongada, puede seguir terminando el clúster, pero primero debe quitar de forma explícita la protección de terminación de este. Esto ayuda a garantizar que EC2 las instancias no se cierren por accidente o error. Puede habilitar la protección de terminación al crear un clúster y también cambiar este ajuste en un clúster en ejecución.

Con la protección de terminación habilitada, la acción `TerminateJobFlows` de la API de Amazon EMR no funciona. Los usuarios no pueden terminar el clúster con esta API ni con el comando

`terminate-clusters` desde la AWS CLI. La API devuelve un error y la CLI se cierra con un código de devolución distinto de cero. Cuando utilice la consola de Amazon EMR para terminar un clúster, se le pedirá que realice un paso adicional para desactivar la protección de terminación.

> **Warning**
>
> La protección de terminación no garantiza que los datos se conserven en caso de un error humano o de una solución alternativa, por ejemplo, si se emite un comando de reinicio desde la línea de comandos mientras se está conectado a la instancia mediante SSH, si una aplicación o un script que se ejecuta en la instancia emite un comando de reinicio o si se utiliza la API de Amazon o EC2 Amazon EMR para deshabilitar la protección de terminación. Esto también es cierto si está ejecutando la versión 7.1 y posteriores de Amazon EMR y una instancia deja de estar en buen estado y es irrecuperable. Incluso con la protección contra la terminación habilitada, los datos guardados en el almacenamiento de instancias, incluidos los datos del HDFS, se pueden perder. Escriba salidas de datos en las ubicaciones de Amazon S3 y cree estrategias de copia de seguridad adecuadas para sus requisitos de continuidad empresarial.

La protección de terminación no afecta a la capacidad para escalar los recursos del clúster con cualquiera de las siguientes acciones:

- Cambiar el tamaño de un clúster manualmente con la tecla O. AWS Management Console AWS CLI Para obtener más información, consulte Cambio manual del tamaño de un clúster de Amazon EMR en ejecución.
- Eliminar instancias de un grupo de instancias secundarias o de tareas mediante una política de escalado descendente con escalado automático. Para obtener más información, consulte Uso del escalado automático con una política personalizada para grupos de instancias en Amazon EMR.
- Eliminar instancias de una flota de instancias reduciendo la capacidad de destino. Para obtener más información, consulte Opciones de flota de instancias.

Protección de rescisión y Amazon EC2

La configuración de protección de terminación de un clúster de Amazon EMR se corresponde con el `DisableApiTermination` atributo de todas las EC2 instancias de Amazon del clúster. Por ejemplo, si habilita la protección de terminación en un clúster de EMR, Amazon EMR se establece automáticamente en true `DisableApiTermination` para todas las instancias EC2 del clúster de

EMR. Lo mismo se aplica si deshabilita la protección de finalización. Amazon EMR se establece automáticamente en false `DisableApiTermination` para todas las EC2 instancias del clúster de EMR. Si cancelas o reduces la escala de un clúster de Amazon EMR y la EC2 configuración de Amazon entra en conflicto en una EC2 instancia, Amazon EMR prioriza la configuración de Amazon EMR sobre la `DisableApiTermination` configuración `DisableApiStop` and de Amazon EC2 y continúa cancelando la instancia. EC2

Por ejemplo, puedes usar la EC2 consola de Amazon para habilitar la protección de terminación en una EC2 instancia de Amazon en un clúster de EMR con la protección de terminación deshabilitada. Si finaliza o reduce el clúster con la consola de Amazon EMR, la o la API de Amazon EMR AWS CLI, Amazon EMR anula la `DisableApiTermination` configuración, la establece en false y termina la instancia junto con las demás instancias.

También puedes usar la EC2 consola de Amazon para habilitar la protección contra interrupciones en una EC2 instancia de Amazon en un clúster de EMR con la protección de terminación deshabilitada. Si finaliza o reduce el clúster, Amazon EMR se establece `DisableApiStop` en false en Amazon EC2 y termina la instancia junto con las demás instancias.

Amazon EMR anula la configuración de `DisableApiStop` solo cuando se finaliza o se reduce verticalmente un clúster. Al activar o desactivar la protección de terminación en un clúster de EMR, Amazon EMR no cambia la `disableApiStop` configuración de ninguna de las EC2 instancias del clúster de EMR correspondiente.

> **Important**
>
> Si crea una instancia como parte de un clúster de Amazon EMR con protección de terminación y utiliza la EC2 API o los AWS CLI comandos de Amazon para modificar la instancia de forma que `DisableApiTermination` sea así y`false`, a continuación, la EC2 API o los AWS CLI comandos de Amazon ejecutan la `TerminateInstances` operación, la EC2 instancia de Amazon finaliza.

Protección de terminación y nodos de YARN en mal estado

Amazon EMR comprueba periódicamente el estado de Apache Hadoop YARN de los nodos que se ejecutan en las EC2 instancias principales y de tareas de Amazon de un clúster. El servicio de verificación de estado informa del estado de NodeManager salud. Si un nodo informa de que su estado es `UNHEALTHY`, el controlador de instancias de Amazon EMR lo agrega a la lista de

denegación y no le asigna contenedores YARN hasta que vuelva a estar en buen estado. Según los estados de la protección de finalización, el reemplazo de nodos en mal estado y la versión de lanzamiento de Amazon EMR, Amazon EMR reemplazará la instancia en mal estado o dejará de asignar controladores a la instancia.

Protección de finalización y finalización tras la ejecución de pasos

Cuando habilita la finalización tras la ejecución de pasos y también habilita la protección de finalización, Amazon EMR ignora la protección de finalización.

Al enviar pasos a un clúster, puede establecer la propiedad `ActionOnFailure` para determinar qué sucede si el paso no puede completar su ejecución debido a un error. Los valores posibles para esta propiedad son `TERMINATE_CLUSTER` (`TERMINATE_JOB_FLOW` con versiones anteriores), `CANCEL_AND_WAIT` y `CONTINUE`. Para obtener más información, consulte Envío del trabajo a un clúster de Amazon EMR.

Si se produce un error en un paso que se ha configurado con el valor `ActionOnFailure` en `CANCEL_AND_WAIT`, y la ejecución de pasos se ha habilitado, el clúster finaliza sin ejecutar los pasos siguientes.

Si se produce un error en un paso que tiene la propiedad `ActionOnFailure` establecida en `TERMINATE_CLUSTER`, utilice la tabla de configuración mostrada a continuación para determinar el resultado.

ActionOnFailure	Finalización después de la ejecución de pasos	Protección de finalización	Resultado
`TERMINATE_CLUSTER`	Habilitado	Deshabilitado	El clúster termina
	Habilitado	Habilitado	El clúster termina
	Deshabilitado	Habilitado	El clúster continúa
	Deshabilitad	Deshabilitado	El clúster termina

Protección de terminación e instancias de spot

La protección de rescisión de Amazon EMR no impide que una instancia EC2 spot de Amazon se cancele cuando el precio spot supera el precio spot máximo.

Configuración de la protección de terminación al lanzar un clúster

Puede activar o desactivar la protección de terminación al lanzar un clúster mediante la consola AWS CLI, la o la API.

Para los clústeres de un solo nodo, la configuración de protección de finalización predeterminada es la siguiente:

- Lanzamiento de un clúster mediante la consola Amazon EMR: la protección de finalización está deshabilitada de forma predeterminada.
- El lanzamiento de un clúster mediante AWS CLI `aws emr create-cluster`: Termination Protection está deshabilitado a menos que `--termination-protected` se especifique lo contrario.
- Lanzamiento de un clúster mediante el RunJobFlowcomando API Amazon EMR: la protección de terminación está deshabilitada a menos que el valor `TerminationProtected` booleano esté establecido en. `true`

Para los clústeres de alta disponibilidad, la configuración de protección de finalización predeterminada es la siguiente:

- Lanzamiento de un clúster mediante la consola de Amazon EMR: la protección de finalización está habilitada de forma predeterminada.
- El lanzamiento de un clúster mediante AWS CLI **`aws emr create-cluster`**: la protección de terminación está deshabilitada a menos que se especifique lo contrario. `--termination-protected`
- Lanzamiento de un clúster mediante el RunJobFlowcomando API Amazon EMR: la protección de terminación está deshabilitada a menos que el valor `TerminationProtected` booleano esté establecido en. `true`

Console

Para activar o desactivar la protección de finalización al crear un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En el caso de la versión de lanzamiento de EMR, elija emr-6.6.0 o posterior.
4. En Finalización del clúster y sustitución del nodo, asegúrese de que la opción Usar la protección de finalización esté preseleccionada o anule la selección para desactivarla.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para activar o desactivar la protección de terminación al crear un clúster mediante el AWS CLI

- Con el AWS CLI, puede lanzar un clúster con la protección de terminación habilitada con el `create-cluster` comando del `--termination-protected` parámetro. La protección de terminación está deshabilitada de forma predeterminada.

  En el siguiente ejemplo, se crea un clúster con la protección de terminación habilitada:

  > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "TerminationProtectedCluster" --release-
label emr-7.10.0 \
--applications Name=Hadoop Name=Hive Name=Pig \
--use-default-roles --ec2-attributes KeyName=myKey --instance-type m5.xlarge \
--instance-count 3 --termination-protected
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Configuración de la protección de terminación para ejecutar clústeres

Puede configurar la protección de terminación para un clúster en ejecución utilizando la consola o la AWS CLI.

Console

Para activar o desactivar la protección de finalización de un clúster en ejecución con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. En la pestaña Propiedades de la página de detalles del clúster, busque Terminación del clúster y seleccione Editar.
4. Seleccione o anule la selección de la casilla Use la protección contra la terminación para activar o desactivar la característica. A continuación, seleccione Guardar cambios para confirmar.

AWS CLI

Para activar o desactivar la protección de terminación para un clúster en ejecución mediante el AWS CLI

- Para habilitar la protección de terminación en un clúster en ejecución mediante la AWS CLI, utilice el comando `modify-cluster-attributes` con el parámetro `--termination-protected`. Para deshabilitarla, utilice el parámetro `--no-termination-protected`.

  El siguiente ejemplo habilita la protección de terminación en el clúster con un ID*`j-3KVTXXXXXX7UG`*:

  ```
  aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --termination-protected
  ```

En el siguiente ejemplo, se deshabilita la protección de terminación en el mismo clúster:

```
aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-termination-
protected
```

Sustitución de nodos en mal estado con Amazon EMR

Amazon EMR utiliza periódicamente el servicio de NodeManager comprobación de estado de Apache Hadoop para supervisar los estados de los nodos principales de sus clústeres de Amazon EMR en Amazon. EC2 Si un nodo no funciona de manera óptima, se marca como en mal estado y el comprobador de estado lo notifica al controlador de Amazon EMR. El controlador Amazon EMR añade el nodo a una lista de denegaciones, lo que impide que el nodo reciba nuevas aplicaciones YARN hasta que el estado del nodo mejore.

> **Note**
>
> Una razón común por la que un nodo está en mal estado es que no tiene espacio en disco. Para obtener más información acerca de cuándo un nodo principal está casi sin espacio en disco, puede resultarle útil el siguiente artículo del Re:post Knowledge Center: ¿Por qué el nodo principal de mi clúster de Amazon EMR se está quedando sin espacio en disco?

> **Note**
>
> Hadoop ofrece la posibilidad de ejecutar comprobaciones personalizadas del estado de los nodos. Esto se explica con más detalle en la documentación de Apache Hadoop en. NodeManager

Puede elegir si Amazon EMR debe finalizar los nodos en mal estado o mantenerlos en el clúster. Si desactiva la sustitución de nodos que no estén en buen estado, permanecerán en la lista de rechazados y seguirán siendo parte de la capacidad del clúster. Aún puedes conectarte a tu instancia EC2 principal de Amazon para la configuración y la recuperación, de modo que puedes cambiar el tamaño del clúster si quieres añadir capacidad. Para obtener más información sobre cómo funcionan la sustitución y la terminación de nodos, consulta Cómo usar la protección de terminación.

Si se activa el reemplazo de nodos en mal estado, Amazon EMR termina un nodo principal en mal estado y aprovisiona una nueva instancia, en función del número de instancias del grupo de instancias o en función de la capacidad objetivo, por ejemplo, las flotas de instancias. Si algún nodo está en mal estado durante más de 45 minutos, Amazon EMR lo sustituirá sin problemas. Si el desmantelamiento correcto de un nodo no se completa en una hora, el nodo se cierra forzosamente, a menos que al terminarlo el clúster quede por debajo del factor de replicación o de las restricciones de capacidad de HDFS.

> **Important**
>
> Tenga en cuenta que el tiempo que tarda un nodo en ser desmantelado o cerrado sin problemas puede estar sujeto a cambios.
> Si bien la sustitución de nodos en mal estado reduce considerablemente la posibilidad de pérdida de datos, no elimina el riesgo por completo. Los datos de HDFS pueden perderse de forma permanente si se reemplaza correctamente una instancia principal en mal estado. Le recomendamos que siempre haga una copia de seguridad de sus datos.

Para obtener más información sobre cómo identificar y recuperar los nodos en mal estado, consulte Errores de recursos. Además, para obtener más información sobre las prácticas recomendadas que puede seguir para mantener el estado de un clúster, consulte la siguiente documentación para conocer el error de recurso El clúster de Amazon EMR termina con NO_SLAVE_LEFT y los nodos principales FAILED_BY_MASTER.

Amazon EMR publica Amazon CloudWatch Events para reemplazar nodos en mal estado, de modo que pueda realizar un seguimiento de lo que sucede con sus instancias principales en mal estado. Para obtener más información, consulte los eventos de sustitución de nodosen mal estado.

Configuración predeterminada de protección de reemplazo y finalización de nodos

La sustitución de nodos en mal estado está disponible en todas las versiones de Amazon EMR, pero la configuración predeterminada depende de la etiqueta de versión que elija. Puede cambiar cualquiera de estos ajustes configurando la sustitución de nodos en mal estado al crear un nuevo clúster o pasando a la configuración del clúster en cualquier momento.

Si va a crear un clúster de un solo nodo o un clúster de alta disponibilidad que ejecute Amazon EMR en su versión 7.0 o anteriores, la configuración predeterminada de sustitución de nodos en mal estado depende de la protección de finalización:

- Al habilitar la protección de finalización, se deshabilita la sustitución de nodos en mal estado.
- La deshabilitación de la protección de finalización permite sustituir los nodos en mal estado.

Configuración de la sustitución de nodos en mal estado al lanzar un clúster

Puede activar o desactivar el reemplazo de nodos en mal estado al lanzar un clúster mediante la consola AWS CLI, la API o la consola.

La configuración predeterminada de sustitución de nodos en mal estado depende de cómo se lance el clúster:

- Consola Amazon EMR: la sustitución de nodos en mal estado está habilitada de forma predeterminada.
- AWS CLI `aws emr create-cluster`— la sustitución de nodos en mal estado está habilitada de forma predeterminada, a menos que lo especifiques`--no-unhealthy-node-replacement`.
- Comando de la RunJobFlow API Amazon EMR: la sustitución de nodos en mal estado está habilitada de forma predeterminada, a menos que establezca el valor `UnhealthyNodeReplacement` booleano en o. `True` `False`

Console

Para activar o desactivar la sustitución de nodos en mal estado al crear un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. Para la versión de lanzamiento de EMR, elija la etiqueta de versión de Amazon EMR que desee.
4. En Finalización del clúster y sustitución del nodo, asegúrese de que la opción Sustitución de nodo en mal estado (recomendada) esté preseleccionada o anule la selección para desactivarla.

5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para activar o desactivar la sustitución de nodos en mal estado al crear un clúster mediante AWS CLI

- Con el AWS CLI, puede lanzar un clúster con el reemplazo de nodos en mal estado habilitado con el `create-cluster` comando del `--unhealthy-node-replacement` parámetro. La sustitución de nodos en mal estado está activada de forma predeterminada.

  En el siguiente ejemplo, se crea un clúster con la opción de sustitución de nodos en mal estado habilitada:

  > **Note**
  >
  > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

  ```
  aws emr create-cluster --name "SampleCluster" --release-label emr-7.10.0 \
  --applications Name=Hadoop Name=Hive Name=Pig \
  --use-default-roles --ec2-attributes KeyName=myKey --instance-type m5.xlarge \
  --instance-count 3 --unhealthy-node-replacement
  ```

  Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte Comandos de Amazon EMR. AWS CLI

Configuración de la sustitución de nodos en mal estado en un clúster en ejecución

Puede activar o desactivar el reemplazo de nodos en mal estado en un clúster en ejecución mediante la consola AWS CLI, la API o la misma.

Console

Para activar o desactivar la sustitución de nodos en mal estado al crear un clúster en ejecución con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. En la pestaña Propiedades de la página de detalles del clúster, busque Finalización del clúster y sustitución del nodo y seleccione Editar.
4. Seleccione o anule la selección de la casilla de verificación de sustitución de nodos en mal estado para activar o desactivar la característica. A continuación, seleccione Guardar cambios para confirmar.

AWS CLI

Para activar o desactivar la sustitución de nodos en mal estado en un clúster en ejecución mediante el AWS CLI

- Para activar la sustitución de nodos en mal estado en un clúster en ejecución con la AWS CLI, utilice el comando `modify-cluster-attributes` con el parámetro `--unhealthy-node-replacement`. Para deshabilitarla, utilice el parámetro `--no-unhealthy-node-replacement`.

  En el siguiente ejemplo, se activa la sustitución de nodos en mal estado en el clúster por un ID*`j-3KVTXXXXXX7UG`*:

  ```
  aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --unhealthy-node-replacement
  ```

  En el siguiente ejemplo, se desactiva la sustitución de nodos en mal estado en el mismo clúster:

  ```
  aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-unhealthy-node-replacement
  ```

Uso de Amazon Linux AMIs en Amazon EMR

Imágenes de Amazon Machine de Amazon Linux (AMIs)

Amazon EMR utiliza una Amazon Machine Image (AMI) de Amazon Linux para inicializar las instancias de EC2 Amazon al crear y lanzar un clúster. La AMI contiene el sistema operativo Amazon Linux, además del software y las configuraciones necesarias para que cada instancia aloje las aplicaciones del clúster.

De forma predeterminada, al crear un clúster, Amazon EMR utiliza una AMI de Amazon Linux (AL) predeterminada que se crea específicamente para la versión de lanzamiento de Amazon EMR que utilice. Para más información sobre la AMI predeterminada de Amazon Linux, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR. Si utiliza la versión 5.7.0 o superior de Amazon EMR, puede optar por especificar una AMI de Amazon Linux personalizada en lugar de la AMI de Amazon Linux predeterminada para Amazon EMR. Una AMI personalizada le permite cifrar el volumen de dispositivo raíz y personalizar las aplicaciones y las configuraciones como alternativa al uso de acciones de arranque. Puede especificar una AMI personalizada para cada tipo de instancia en la configuración del grupo de instancias o de la flota de instancias de un clúster de Amazon EMR. La compatibilidad con varias AMI personalizadas le brinda la flexibilidad de utilizar más de un tipo de arquitectura en un clúster. Consulte Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR.

Amazon EMR adjunta automáticamente un volumen SSD de uso general de Amazon EBS como dispositivo raíz para todos. AMIs Mejoran el rendimiento con el respaldo de EBS AMIs . Para obtener más información sobre Amazon Linux AMIs, consulte Amazon Machine Images (AMI). Para más información sobre el almacenamiento de instancias para las instancias de Amazon EMR, consulte Comportamiento y opciones de almacenamiento de instancias en Amazon EMR.

Temas

- Uso de la AMI de Amazon Linux predeterminada para Amazon EMR
- Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR
- Cambio de versión de Amazon Linux al crear un clúster de EMR
- Personalización del volumen de dispositivo raíz de Amazon EBS

Uso de la AMI de Amazon Linux predeterminada para Amazon EMR

Cada versión de lanzamiento de Amazon EMR utiliza una AMI de Amazon Linux predeterminada para Amazon EMR a menos que se especifique una AMI personalizada. A partir de Amazon EMR 5.36, Amazon EMR 6.6 y Amazon EMR 7.0, el comportamiento predeterminado para actualizar Amazon Linux 2 (para EMR 5.x y 6.x, 023 AL2 para AL2 EMR 7.x) en una AMI predeterminada de Amazon EMR es aplicar automáticamente la última versión de Amazon Linux a la AMI EMR predeterminada de Amazon.

Actualizaciones automáticas de Amazon Linux para las versiones de Amazon EMR

Al lanzar un clúster con la versión de revisión más reciente de Amazon EMR 7.0, 6.6, 5.36 o sus respectivas versiones posteriores, Amazon EMR utiliza la versión más reciente de Amazon Linux para la AMI de Amazon EMR predeterminada. Por ejemplo:

- Cuando hay una versión `x.x.0` y una versión `x.x.1`, la versión `x.x.0` deja de recibir actualizaciones de la AMI cuando se lanza `x.x.1`.
- Del mismo modo, `x.x.1` deja de recibir actualizaciones de la AMI cuando se lanza `x.x.2`.
- Más adelante, cuando se lanza `x.y.0`, `x.x.[latest]` sigue recibiendo actualizaciones de la AMI al mismo tiempo que `x.y.[latest]`.

Para comprobar si está utilizando la versión de revisión más reciente indicada por el número que aparece después del segundo punto decimal (`6.8.1`) de una versión de Amazon EMR, consulte las versiones disponibles en la Guía de versiones de Amazon EMR, consulte el menú desplegable Versión de Amazon EMR al crear un clúster en la consola o utilice la API `ListReleaseLabels` o la acción `list-release-labels` de la CLI. Para recibir actualizaciones cuando lancemos una nueva versión de Amazon EMR, suscríbase a la fuente RSS de la página Novedades de la Guía de versiones.

Si lo desea, puede optar por lanzar el clúster con la versión de Amazon Linux con la que se envió por primera vez la versión de Amazon EMR. Para obtener información sobre cómo especificar la versión de Amazon Linux para su clúster, consulte Cambio de versión de Amazon Linux al crear un clúster de EMR.

Versiones de Amazon Linux predeterminadas

Temas

- Predeterminado AMIs para Amazon EMR 7.0 y versiones posteriores

- Predeterminado AMIs para Amazon EMR 6.6 y versiones posteriores
- Predeterminado AMIs para Amazon EMR 5.x

Predeterminado AMIs para Amazon EMR 7.0 y versiones posteriores

En la siguiente tabla se muestra la información de Amazon Linux sobre la versión de parche más reciente de las versiones 7.0 y posteriores de Amazon EMR.

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.8.2 250721.2	6.1.144-170.251.am zn2023	14 de agosto de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 • eu-north-1 • eu-central-2 • ap-east-1 • ap-south-2 • ap-southeast-3 • ap-southeast-5 • ap-southeast-4 • ap-south-1 • ap-northeast-3 • ap-northeast-2

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ap-southeast-1 • ap-southeast-2 • ap-southe ast-7 (7.1.0+) • ap-northeast-1 • af-south-1 • sa-east-1 • me-south-1 • me-central-1 • ca-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northwest-1 • il-central-1 • mx-centra l-1 (7.1.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.7.2 250623,1	6.1.141-155.222. amzn2023	21 de julio de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 • eu-north-1 • eu-central-2 • ap-east-1 • ap-south-2 • ap-southeast-3 • ap-southeast-5 • ap-southeast-4 • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2 • ap-southe ast-7 (7.1.0+) • ap-northeast-1 • af-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• sa-east-1 • me-south-1 • me-central-1 • ca-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northwest-1 • il-central-1 • mx-centra l-1 (7.1.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023,7,2 250609,(	6.1.140 - 154.222. a.m. de 2023	14 de julio de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 • eu-north-1 • eu-central-2 • ap-east-1 • ap-south-2 • ap-southeast-3 • ap-southeast-5 • ap-southeast-4 • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2 • ap-southe ast-7 (7.1.0+) • ap-northeast-1 • af-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• sa-east-1 • me-south-1 • me-central-1 • ca-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northwest-1 • il-central-1 • mx-centra l-1 (7.1.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.7.2 250527,1	6.1.134 - 152.225. a.m. de 2023	19 de junio de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 • eu-north-1 • eu-central-2 • ap-east-1 • ap-south-2 • ap-southeast-3 • ap-southeast-5 • ap-southeast-4 • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2 • ap-southe ast-7 (7.1.0+) • ap-northeast-1 • af-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• sa-east-1 • me-south-1 • me-central-1 • ca-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northwest-1 • il-central-1 • mx-centra l-1 (7.1.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.7.2 250512,(	6.1.134 - 152.225. a.m. de 2023	4 de junio de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 • eu-north-1 • eu-central-2 • ap-east-1 • ap-south-2 • ap-southeast-3 • ap-southeast-5 • ap-southeast-4 • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2 • ap-southe ast-7 (7.1.0+) • ap-northeast-1 • af-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• sa-east-1 • me-south-1 • me-central-1 • ca-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northwest-1 • il-central-1 • mx-centra l-1 (7.1.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.7.2 250428,	6.1.134-150.224. a.m. de 2023	23 de mayo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 • eu-north-1 • eu-central-2 • ap-east-1 • ap-south-2 • ap-southeast-3 • ap-southeast-5 • ap-southeast-4 • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2 • ap-southe ast-7 (7.1.0+) • ap-northeast-1 • af-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• sa-east-1 • me-south-1 • me-central-1 • ca-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northwest-1 • il-central-1 • mx-centra l-1 (7.1.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.7.2 250414,(	6.1.132-147.221. a.m. de 2023	12 de mayo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • ap-southeast-7 • ap-southeast-5 • mx-central-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.7.2 250331.(	6.1.131-143,221. amzn2023	18 de abril de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • ap-southeast-7 • ap-southeast-5 • mx-central-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.6.2 250303.(	6.1.129-138.220. amzn2023	27 de marzo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • ap-southeast-7 • ap-southeast-5 • mx-central-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.6.2 250218,2	6.1.128-136.201. amzn2023	08 de marzo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • ap-southeast-7 • ap-southeast-5 • mx-central-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.6.2 250211,(	6.1.127-135.201. amzn2023	26 de febrero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • ap-southeast-7 • ap-southeast-5 • mx-central-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.6.2 250123.4	6.1.124-134.200. amzn2023	27 de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.6.2 250115.(	6.1.119-129.201. amzn2023	23 de enero de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.6.2 241121,(	6.1.115-126.197. a.m. de 2023	12 de diciembre de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.5.2 241031,(	6.1.112-124.190. amzn2023	15 de noviembre de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2023.5.2 241001.1	6.1.109-118.189.am zn2023	4 de octubre de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20235,2(40819,0	6.1.102-111.182.am zn2023	20 de agosto de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20235,20 40730,00	6.1.97-104.177.amzn2023	2 de agosto de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20235,20 40722,0	6.1.97-104.177.amzn2023	24 de julio de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20235,20 40708,0	6.1.96-102.177.amzn2023	23 de julio de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20233,2(40304,0	6.1.79-99.164.amzn2023	12 de marzo de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20233,20 40219,0	6.1.77-99.164.amzn2023	1 de marzo de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20233,2(40205,0	6.1.75-99.163.amzn2023	19 de febrero de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20233,20 40122,0	6.1.72-96.166.amzn2023	5 de febrero de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20233,2(40108,0	6.1.72-96.166.amzn2023	24 de enero de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • ca-west-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
20233,20 31211,4	6.1.66-91.160.amzn2023	19 de diciembre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• ca-central-1 • il-central-1 • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

Predeterminado AMIs para Amazon EMR 6.6 y versiones posteriores

En la siguiente tabla, se muestra la información de Amazon Linux sobre la versión de revisión más reciente a partir de las versiones 6.6.x de Amazon EMR.

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 610.0	4.14.355-277.647.amzn2	14 de julio de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 (6,7,0 +) • eu-north-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• eu-central-2 (6,7,0 +) • ap-east-1 • ap-south-2 (6,7,0 +) • ap-southeast-3 • ap-southeast-5 (6.11.1 y 6.14.0+) • ap-southeast-4 (6,80+) • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2 • ap-southeast-7 (6.11.1 y 6.14.0+) • ap-northeast-1 • af-south-1 • sa-east-1 • me-south-1 • me-central-1 (6.7.0+) • ca-central-1 • ca-west-1 (6,9,0 +) • us-gov-east-1 • us-gov-west-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `cn-north-1` • `cn-northwest-1` • `il-central-1` (6,7,0+) • `mx-central-1` (6.11.1 y 6.14.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 527,1	4.14.355-277.647.amzn2	19 de junio de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 (6,7,0 +) • eu-north-1 • eu-central-2 (6,7,0 +) • ap-east-1 • ap-south-2 (6,7,0 +) • ap-southeast-3 • ap-southe ast-5 (6.11.1 y 6.14.0+) • ap-southe ast-4 (6,80+) • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ap-southeast-7` (6.11.1 y 6.14.0+) • `ap-northeast-1` • `af-south-1` • `sa-east-1` • `me-south-1` • `me-central-1` (6.7.0+) • `ca-central-1` • `ca-west-1` (6,9,0 +) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northwest-1` • `il-central-1` (6,7,0 +) • `mx-central-1` (6.11.1 y 6.14.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 512,0	4.14.355-277.643.amzn2	4 de junio de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 (6,7,0 +) • eu-north-1 • eu-central-2 (6,7,0 +) • ap-east-1 • ap-south-2 (6,7,0 +) • ap-southeast-3 • ap-southeast-5 (6.11.1 y 6.14.0+) • ap-southeast-4 (6.8.0+) • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ap-southeast-7` (6.11.1 y 6.14.0+) • `ap-northeast-1` • `af-south-1` • `sa-east-1` • `me-south-1` • `me-central-1` (6.7.0+) • `ca-central-1` • `ca-west-1` (6,9,0 +) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northwest-1` • `il-central-1` (6,7,0 +) • `mx-central-1` (6.11.1 y 6.14.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 428,0	4.14.355-276.639.amzn2	23 de mayo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 (6.7,0+) • eu-north-1 • eu-central-2 (6,7,0 +) • ap-east-1 • ap-south-2 (6,7,0 +) • ap-southeast-3 • ap-southe ast-5 (6.11.1 y 6.14.0+) • ap-southe ast-4 (6.8.0+) • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ap-southeast-7` (6.11.1 y 6.14.0+) • `ap-northeast-1` • `af-south-1` • `sa-east-1` • `me-south-1` • `me-central-1` (6.7.0+) • `ca-central-1` • `ca-west-1` (6,9,0 +) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northwest-1` • `il-central-1` (6,7,0 +) • `mx-central-1` (6.11.1 y 6.14.0+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 414,0	4.14.355-276.618.amzn2	12 de mayo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 321.0	4,14,355	09 de abril de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 305.0	4,14,355	18 de marzo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 220,0	4,14,355	8 de marzo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 201,0	4,14,355	28 de febrero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 123.4	4,14,355	27 de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 116.0	4,14,355	23 de enero de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 217,0	4,114,355	8 de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 001.0	4,14,352	4 de octubre de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 816,0	4,14,350	21 de agosto de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 809.0	4,14,349	20 de agosto de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 719,0	4,14,348	25 de julio de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 709,1	4.14.348	23 de julio de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 223.0	4,14,336	8 de marzo de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 131,0	4,14,336	14 de febrero de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 124.0	4,14.336	7 de febrero de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 109,0	4,1334	24 de enero de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 218,0	4,14,330	2 de enero de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 206.0	4,14.330	22 de diciembre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 116,0	4,14,328	11 de diciembre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 101.0	4.14.327	17 de noviembre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 020,1	4,14326	07 de noviembre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 012,1	4,14326	26 de octubre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 926,0	4,14,322	19 de octubre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.8+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 8906.0	4,14,322	04 de octubre de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.9+ y 5.36.1)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 822,0	4,14,322	30 de agosto de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.9+ y 5.36.1)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 808.0	4,14320	24 de agosto de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-central-1 (6.10+) • me-south-1 • ca-central-1 • il-central-1 (6.9+ y 5.36.1) • us-gov-east-1 • us-gov-west-1 • cn-north-1 • cn-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 727,0	4,14320	14 de agosto de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.9+ y 5.36.1)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 719,0	4,14320	2 de agosto de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-southeast-4 (6.8+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.9+ y 5.36.1)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 628,0	4,14318	12 de julio de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 (6.10+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 612,0	4.14.314	23 de junio de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 (6.10+)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 504.1	4,14.313	16 de mayo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10+) • eu-south-1 • eu-south-2 (6.10+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10+) • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 418,0	4,14,311	3 de mayo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (solo 6.10) • eu-south-1 • eu-south-2 (solo 6.10) • ap-east-1 • ap-south-1 • ap-south-2 (solo 6.10) • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 404.1	4,14,311	18 de abril de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 404.0	4,14,311	10 de abril de 2023	• `us-east-1` • `eu-west-3`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 320,0	4,14309	30 de marzo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 307.0	4,14.305	15 de marzo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 207,0	4,14,304	3 de marzo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 • ap-east-1 • ap-south-1 • ap-south-2 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 119,1	4,14.301	9 de febrero de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 210.1	4.14.301	12 de enero de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 912,1	4,14,291	7 de octubre de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1
2.0.2022 805.0	4,14,287	30 de agosto de 2022	• us-west-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 719,0	4,14,287	10 de agosto de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 426,0	4,14,281	10 de junio de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 406.1	4,14,275	2 de mayo de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

Predeterminado AMIs para Amazon EMR 5.x

En la siguiente tabla se muestra la información de Amazon Linux sobre la versión de revisión más reciente de Amazon EMR 5.x, versiones 5.36 y posteriores.

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 428.0	4.14.355-276.639.amzn2	23 de mayo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-central-1 • eu-west-1 • eu-west-2 • eu-south-1 • eu-west-3 • eu-south-2 • eu-north-1 • eu-central-2 • ap-east-1 • ap-south-2 • ap-southeast-3 • ap-southe ast-5 (5.36.2) • ap-southeast-4 • ap-south-1 • ap-northeast-3 • ap-northeast-2 • ap-southeast-1 • ap-southeast-2 • ap-southe ast-7 (5.36.2) • ap-northeast-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• af-south-1 • sa-east-1 • me-south-1 • me-central-1 • ca-central-1 • ca-west-1 (5,36,1 +) • us-gov-east-1 (5,36,1 +) • us-gov-west-1 (5,36,1 +) • cn-north-1 (5,36,1 +) • cn-northwest-1 (5,36,1 +) • il-central-1 • mx-central-1 (5.36.2)

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 321.0	4,14,355	09 de abril de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 305.0	4,14,355	18 de marzo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 220,0	4,14,355	8 de marzo de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 201,0	4,14,355	28 de febrero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 123.4	4,14,355	27 de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2025 116.0	4,114,355	23 de enero de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 217,0	4,114,355	8 de enero de 2025	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-central-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `ca-central-1` • `il-central-1` • `ca-west-1` (6.9.0+ y 5.36.1+) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 816,0	4,14,350	21 de agosto de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 809.0	4,14,349	20 de agosto de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 719,0	4,14,348	25 de julio de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2024 709,1	4.14.348	23 de julio de 2024	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-central-2 (6.10.1+) • eu-south-1 • eu-south-2 (6.10.1+) • ap-east-1 • ap-south-1 • ap-south-2 (6.10.1+) • ap-southeast-3 • ap-southeast-4 (6.8.1+ y 5.36.1) • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
			• `me-central-1` (6.10.1+) • `me-south-1` • `ca-central-1` • `il-central-1` (6.8.1+ y 5.36.1) • `ca-west-1` (6.9.1+ y 5.36.1) • `us-gov-east-1` • `us-gov-west-1` • `cn-north-1` • `cn-northeast-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 504.1	4,14.313	16 de mayo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • ca-central-1 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • me-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 418,0	4,14,311	3 de mayo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • ca-central-1 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • me-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 404.1	4,14,311	18 de abril de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • ca-central-1 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1
2.0.2023 404.0	4,14,311	10 de abril de 2023	• us-east-1 • eu-west-3

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 320,0	4,14309	30 de marzo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • ca-central-1 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 307.0	4,14.305	15 de marzo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • ca-central-1 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2023 207,0	4,14,304	3 de marzo de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 210.1	4.14.301	12 de enero de 2023	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 103.3	4,14.296	5 de diciembre de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 004.0	4,14,294	2 de noviembre de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 912,1	4,14,291	7 de octubre de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 719,0	4,14,287	10 de agosto de 2022	• `us-west-1` • `eu-west-3` • `eu-north-1` • `eu-central-1` • `ap-south-1` • `me-south-1`

OsRelea Label (versión AL)	Versión del kernel de AL	Fecha de disponibilidad	Regiones de AWS
2.0.2022 426,0	4,14,281	14 de junio de 2022	• us-east-1 • us-east-2 • us-west-1 • us-west-2 • eu-north-1 • eu-west-1 • eu-west-2 • eu-west-3 • eu-central-1 • eu-south-1 • ap-east-1 • ap-south-1 • ap-southeast-3 • ap-northeast-1 • ap-northeast-2 • ap-northeast-3 • ap-southeast-1 • ap-southeast-2 • af-south-1 • sa-east-1 • me-south-1 • ca-central-1

Consideraciones sobre las actualizaciones de software

Tome nota de los siguientes comportamientos predeterminados de actualización de software:

Amazon EMR 7.x: Amazon Linux 2023

Las versiones 7.0 y superiores de Amazon EMR se ejecutan en Amazon Linux 2023 (AL2023). Sus clústeres solo contienen actualizaciones de seguridad que estaban disponibles en la versión de la AMI AL2 023 que eligió al crearlas. En el momento del lanzamiento, las instancias de clúster de EMR no instalarán las actualizaciones de seguridad más recientes de los repositorios de paquetes habilitados. Para recibir las actualizaciones de seguridad más recientes, le recomendamos que vuelva a crear el clúster periódicamente. Para obtener más información sobre AL2 023, consulte Actualización de Amazon Linux 2023 en la Guía del usuario de Amazon Linux 2023.

Amazon EMR 5.36+ y 6.6+ — Amazon Linux 2

Las versiones 5.36 y posteriores de Amazon EMR, y 6.6 y posteriores, se ejecutan en Amazon Linux 2 (). AL2 Sus clústeres solo contienen actualizaciones de seguridad que estaban disponibles en la versión de AL2 AMI que eligió cuando las creó. En el momento del lanzamiento, las instancias de clúster de EMR no instalarán las actualizaciones de seguridad más recientes de los repositorios de paquetes habilitados. Para recibir las actualizaciones de seguridad más recientes, le recomendamos que vuelva a crear el clúster periódicamente. Para obtener más información AL2, consulte Administrar el software de la AL2 instancia en la Guía del usuario de Amazon Linux 2.

Amazon EMR 3.x, 4.x, 5.0.0 a 5.35.0 y 6.0.0 a 6.5.0: la AMI de Amazon Linux está bloqueada en la versión de lanzamiento de Amazon EMR

Para las versiones 3.x, 4.x, 5.0.0 a 5.35.0 y 6.0.0 a 6.5.0 de Amazon EMR, la AMI predeterminada se basa en la mayoría de las AMI de Amazon Linux disponibles en el momento de la publicación de up-to-date Amazon EMR. Se ha probado la compatibilidad de la AMI con las aplicaciones de big data y las funciones de Amazon EMR incluidas en esa versión de lanzamiento.

Cuando una EC2 instancia de Amazon se inicia por primera vez en un clúster basado en la AL2 AMI o Amazon Linux (AL) predeterminada para Amazon EMR, comprueba si hay actualizaciones de software que se apliquen a la versión de lanzamiento en los repositorios de paquetes habilitados para AL y Amazon EMR. Al igual que con otras EC2 instancias de Amazon que ejecutan AL o AL2 AMIs, las actualizaciones de seguridad críticas e importantes de estos repositorios se instalan automáticamente.

Tenga en cuenta también que, en la configuración de red, debe permitir la salida de HTTP y HTTPS a los repositorios de Amazon Linux en Amazon S3. De lo contrario, las actualizaciones de seguridad no se llevarán a cabo correctamente. Para obtener más información, consulte Amazon Linux - Package repository en la Guía del EC2 usuario de Amazon. De forma predeterminada, otros

paquetes de software y actualizaciones del kernel que requieren un reinicio, incluidos NVIDIA y CUDA, no se descargan automáticamente en el primer arranque.

> **Important**
>
> Las versiones 3.x, 4.x, 5.0.0 a 5.35.0 y 6.0.0 a 6.5.0 de Amazon EMR utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieran un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan el AL o predeterminado AL2 AMIs. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan automáticamente esas actualizaciones. Para obtener las actualizaciones del núcleo, puede utilizar la AMI de Amazon Linux más reciente, tal y como se describe en Uso de una AMI personalizada, a fin de ofrecer más flexibilidad a la configuración del clúster de Amazon EMR.

El clúster se lanza con o sin actualizaciones

Tenga en cuenta que, si las actualizaciones de software no se pueden instalar porque no se puede acceder a los repositorios de paquetes la primera vez que se inicia el clúster, la instancia del clúster sigue completando su lanzamiento. Por ejemplo, es posible que no se pueda acceder a los repositorios porque S3 no está disponible temporalmente o puede que tenga reglas de VPC o firewall configuradas para bloquear el acceso.

No se ejecuta **`sudo yum update`**

Cuando se conecta a una instancia de clúster mediante SSH, las primeras líneas de la salida en pantalla proporcionan un enlace a las notas de la versión de la AMI de Amazon Linux que utiliza la instancia, un aviso de la versión más reciente de la AMI de Amazon Linux, un aviso del número de paquetes disponibles para la actualización desde los repositorios habilitados y una directiva para ejecutar `sudo yum update`.

> **Important**
>
> Se recomienda no ejecutar `sudo yum update` en las instancias de clúster, ni mientras está conectado con SSH ni mediante una acción de arranque. Esto puede causar incompatibilidades, ya que todos los paquetes se instalan de forma indiscriminada.

Prácticas recomendadas de actualización de software

Prácticas recomendadas para administrar las actualizaciones de software

- Si utiliza una versión anterior de Amazon EMR, es conveniente que pruebe una migración a la versión más reciente antes de actualizar los paquetes de software.
- Si migra a una versión posterior o actualiza los paquetes de software, pruebe primero la implementación en un entorno que no sea de producción. La opción para clonar clústeres mediante la consola de Amazon EMR resulta muy útil para esto.
- Evalúe las actualizaciones de software para las aplicaciones y para la versión de la AMI de Amazon Linux por separado. Únicamente pruebe e instale en entornos de producción los paquetes que considere absolutamente necesarios para el nivel de seguridad requerido y la funcionalidad o el rendimiento de las aplicaciones.
- Visite el Centro de seguridad de Amazon Linux para comprobar si existen actualizaciones.
- Evite instalar paquetes conectándose a las instancias de clúster individuales mediante SSH. En su lugar, utilice una acción de arranque para instalar y actualizar los paquetes en todas las instancias de clúster que lo necesiten. Para ello, es necesario que termine un clúster y lo vuelva a lanzar. Para obtener más información, consulte Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR.

Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR

Si utiliza la versión 5.7.0 o superior de Amazon EMR, puede optar por especificar una AMI de Amazon Linux personalizada en lugar de la AMI de Amazon Linux predeterminada para Amazon EMR. Una AMI personalizada resulta útil si desea hacer lo siguiente:

- Preinstalar aplicaciones y realizar otras personalizaciones en lugar de utilizar acciones de arranque. Esto puede mejorar el tiempo de inicio del clúster y simplificar el flujo de trabajo de inicio.

Para obtener más información y un ejemplo, consulte Creación de una AMI de Amazon Linux personalizada a partir de una instancia preconfigurada.

- Implementar configuraciones de clúster y de nodo más sofisticadas de lo que permiten las acciones de arranque.
- Cifre los volúmenes de los dispositivos raíz de EBS (volúmenes de arranque) de las EC2 instancias de su clúster si utiliza una versión de Amazon EMR anterior a la 5.24.0. Al igual que con la AMI predeterminada, el tamaño mínimo del volumen raíz de una AMI personalizada es de 10 GiB para las versiones 6.9 y posteriores de Amazon EMR, y de 15 GiB para las versiones 6.10 y posteriores de Amazon EMR. Para obtener más información, consulte Creación de una AMI personalizada con un volumen de dispositivo raíz de Amazon EBS cifrado.

> **Note**
>
> A partir de la versión 5.24.0 de Amazon EMR, puede utilizar una opción de configuración de seguridad para cifrar los volúmenes de almacenamiento y los dispositivos raíz de EBS si lo especifica como proveedor de claves. AWS KMS Para obtener más información, consulte Cifrado de disco local.

Debe existir una AMI personalizada en la misma AWS región en la que se creó el clúster. También debe coincidir con la arquitectura de la EC2 instancia. Por ejemplo, una instancia m5.xlarge tiene una arquitectura x86_64. Por lo tanto, para aprovisionar una instancia m5.xlarge mediante una AMI personalizada, la AMI también debe tener una arquitectura x86_64. Del mismo modo, para aprovisionar una instancia m6g.xlarge, que tiene una arquitectura arm64, la AMI personalizada debe tener una arquitectura arm64. Para obtener más información sobre cómo identificar una AMI de Linux para tu tipo de instancia, consulta Buscar una AMI de Linux en la Guía del EC2 usuario de Amazon.

> **Important**
>
> Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan

automáticamente esas actualizaciones. Para obtener actualizaciones del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.

Creación de una AMI de Amazon Linux personalizada a partir de una instancia preconfigurada

Los pasos básicos para preinstalar el software y realizar otras configuraciones para crear una AMI de Amazon Linux personalizada para Amazon EMR son los siguientes:

- Lance una instancia desde la AMI de Amazon Linux de base.
- Conéctese a la instancia para instalar software y realizar otras personalizaciones.
- Cree una nueva imagen (instantánea de AMI) de la instancia que haya configurado.

Después de crear la imagen en función de su instancia personalizada, puede copiar dicha imagen a un destino cifrado tal y como se describe en Creación de una AMI personalizada con un volumen de dispositivo raíz de Amazon EBS cifrado.

Tutorial: Creación de una AMI desde una instancia con software personalizado Instalado

Para lanzar una EC2 instancia basada en la AMI de Amazon Linux más reciente

1. Use el AWS CLI para ejecutar el siguiente comando, que crea una instancia a partir de una AMI existente. *MyKeyName*Sustitúyalo por el par de claves que utilizas para conectarte a la instancia y *MyAmiId* por el ID de una AMI de Amazon Linux adecuada. Para ver la AMI más reciente IDs, consulte la AMI de Amazon Linux.

   > **Note**
   > Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

   ```
   aws ec2 run-instances --image-id MyAmiID \
   --count 1 --instance-type m5.xlarge \
   ```

```
--key-name MyKeyName --region us-west-2
```

El valor de salida `InstanceId` se utiliza como *`MyInstanceId`* en el siguiente paso.

2. Ejecuta el siguiente comando:

```
aws ec2 describe-instances --instance-ids MyInstanceId
```

El valor de salida `PublicDnsName` se utiliza para conectarse a la instancia en el siguiente paso.

Para conectar a la instancia e instalar software

1. Utilice una conexión SSH que le permita ejecutar comandos de shell en su instancia de Linux. Para obtener más información, consulta Cómo conectarse a tu instancia de Linux mediante SSH en la Guía del EC2 usuario de Amazon.
2. Realice las personalizaciones necesarias. Por ejemplo:

```
sudo yum install MySoftwarePackage
sudo pip install MySoftwarePackage
```

Para crear una instantánea desde su imagen personalizada

- Después de personalizar la instancia, utilice el comando `create-image` para crear una AMI desde la instancia.

```
aws ec2 create-image --no-dry-run --instance-id MyInstanceId --name MyEmrCustomAmi
```

El valor de salida `imageID` se utiliza al lanzar el clúster o crear una instantánea cifrada. Para obtener más información, consulte Utilice una sola AMI personalizada en un clúster de EMR y Creación de una AMI personalizada con un volumen de dispositivo raíz de Amazon EBS cifrado.

Cómo utilizar una AMI personalizada en un clúster de Amazon EMR

Puede utilizar una AMI personalizada para aprovisionar un clúster de Amazon EMR de dos maneras:

- Use una sola AMI personalizada para todas las EC2 instancias del clúster.

- Utilice una personalización diferente AMIs para los distintos tipos de EC2 instancias que se utilizan en el clúster.

Puede usar solo una de las dos opciones al aprovisionar un clúster de EMR y no puede cambiarla una vez que se haya iniciado el clúster.

Consideraciones sobre el uso de la personalización única frente a la personalización múltiple AMIs en un clúster de Amazon EMR

Consideración	AMI personalizada única	Múltiples personalizaciones AMIs
Utilice procesadores x86 y Graviton2 con procesadores personalizados AMIs en el mismo clúster	✕ No compatible	✓ Compatible
La personalización de la AMI varía según el tipo de instancia	✕ No compatible	✓ Compatible
Cambie la personalización AMIs al agregar una nueva instancia de tarea. groups/fleets to a running cluster. Note: you cannot change the custom AMI of existing instance groups/fleets	✕ No compatible	✓ Compatible
Use AWS la consola para iniciar un clúster	✓ Compatible	✕ No compatible
Se utiliza AWS CloudFormation para iniciar un clúster	✓ Compatible	✓ Compatible

Utilice una sola AMI personalizada en un clúster de EMR

Para especificar el ID de una AMI personalizada al crear un clúster, utilice una de las siguientes opciones:

- AWS Management Console
- AWS CLI
- SDK de Amazon EMR
- API de Amazon EMR RunJobFlow
- AWS CloudFormation (consulte la `CustomAmiID` propiedad en Cluster InstanceGroupConfig, Cluster InstanceTypeConfig InstanceGroupConfig, Resource o Resource InstanceFleetConfig -) InstanceTypeConfig

Amazon EMR console

Para especificar una sola AMI personalizada desde la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Nombre y aplicaciones, busque Opciones del sistema operativo. Elija AMI personalizada e ingrese el ID de su AMI en el campo AMI personalizada.
4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para especificar una sola AMI personalizada con el AWS CLI

- Utilice el parámetro `--custom-ami-id` para especificar el ID de AMI al ejecutar el comando `aws emr create-cluster`.

  El siguiente ejemplo especifica un clúster que utiliza una sola AMI personalizada con un volumen de arranque de 20 GiB. Para obtener más información, consulte Personalización del volumen de dispositivo raíz de Amazon EBS.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Cluster with My Custom AMI" \
--custom-ami-id MyAmiID --ebs-root-volume-size 20 \
--release-label emr-5.7.0 --use-default-roles \
--instance-count 2 --instance-type m5.xlarge
```

Utilice varias opciones personalizadas AMIs en un clúster de Amazon EMR

Para crear un clúster con varias opciones personalizadas AMIs, utilice una de las siguientes opciones:

- AWS CLI versión 1.20.21 o superior
- AWS SDK
- Amazon EMR RunJobFlowen la referencia de la API de Amazon EMR
- AWS CloudFormation (consulte la `CustomAmiID` propiedad en Cluster InstanceGroupConfig, InstanceTypeConfigCluster, Resource o InstanceGroupConfigResource InstanceFleetConfig -) InstanceTypeConfig

Actualmente, la consola de AWS administración no admite la creación de un clúster mediante la personalización múltiple AMIs.

Example - Usa la AWS CLI para crear un clúster de grupos de instancias mediante múltiples opciones personalizadas AMIs

Con la versión 1.20.21 o superior de la AWS CLI, puede asignar una sola AMI personalizada a todo el clúster o puede asignar varias personalizadas AMIs a cada nodo de instancia del clúster.

En el siguiente ejemplo, se muestra un clúster de grupos de instancias uniforme creado con dos tipos de instancias (m5.xlarge) que se utilizan en todos los tipos de nodos (principal, central y de

tarea). Cada nodo tiene varias opciones personalizadas. AMIs En el ejemplo se ilustran varias características de la configuración de varias AMI personalizadas:

- No hay ninguna AMI personalizada asignada en el nivel de clúster. Esto es para evitar conflictos entre la AMI personalizada múltiple AMIs y una única personalizada, lo que provocaría un error en el lanzamiento del clúster.
- El clúster puede tener varias configuraciones personalizadas AMIs en los nodos de tareas principales, principales e individuales. Esto permite realizar personalizaciones de AMI individuales, como aplicaciones preinstaladas, configuraciones de clústeres sofisticadas y volúmenes de dispositivos raíz de Amazon EBS cifrados.
- El nodo principal del grupo de instancias solo puede tener un tipo de instancia y la AMI personalizada correspondiente. Del mismo modo, el nodo principal solo puede tener un tipo de instancia y la AMI personalizada correspondiente.
- El clúster puede tener varios nodos de tarea.

```
aws emr create-cluster --instance-groups
InstanceGroupType=PRIMARY,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-123456
InstanceGroupType=CORE,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-234567
InstanceGroupType=TASK,InstanceType=m6g.xlarge,InstanceCount=1,CustomAmiId=ami-345678
InstanceGroupType=TASK,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-456789
```

Example - Usa la versión 1.20.21 o superior de la AWS CLI para agregar un nodo de tareas a un clúster de grupos de instancias en ejecución con varios tipos de instancias y varios personalizados AMIs

Con la versión 1.20.21 o superior de la AWS CLI, puedes añadir varios grupos personalizados AMIs a un grupo de instancias y añadirlos a un clúster en ejecución. El argumento `CustomAmiId` se puede usar con el comando `add-instance-groups`, como se muestra en el siguiente ejemplo. Observe que se utiliza el mismo ID de AMI personalizada múltiple (ami-123456) en más de un nodo.

```
aws emr create-cluster --instance-groups
InstanceGroupType=PRIMARY,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-123456
InstanceGroupType=CORE,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-123456
InstanceGroupType=TASK,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-234567

{
    "ClusterId": "j-123456",
    ...
```

```
}

aws emr add-instance-groups --cluster-id j-123456 --instance-groups
 InstanceGroupType=Task,InstanceType=m6g.xlarge,InstanceCount=1,CustomAmiId=ami-345678
```

Example - Utilice la AWS CLI versión 1.20.21 o superior para crear un clúster de flota de instancias, varios tipos de instancias personalizados AMIs, varios tipos de instancias, principal bajo demanda, núcleo bajo demanda, múltiples nodos de núcleo y tareas

```
aws emr create-cluster --instance-fleets
InstanceFleetType=PRIMARY,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarg
 CustomAmiId=ami-123456}']
InstanceFleetType=CORE,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarge,C
{InstanceType=m6g.xlarge, CustomAmiId=ami-345678}']
InstanceFleetType=TASK,TargetSpotCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarge,Custo
{InstanceType=m6g.xlarge, CustomAmiId=ami-567890}']
```

Example - Utilice la versión 1.20.21 o superior de la AWS CLI para añadir nodos de tareas a un clúster en ejecución con varios tipos de instancias y varios personalizados AMIs

```
aws emr create-cluster --instance-fleets
InstanceFleetType=PRIMARY,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarg
 CustomAmiId=ami-123456}']
InstanceFleetType=CORE,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarge,C
{InstanceType=m6g.xlarge, CustomAmiId=ami-345678}']

{
    "ClusterId": "j-123456",
    ...
}

aws emr add-instance-fleet --cluster-id j-123456 --instance-fleet
InstanceFleetType=TASK,TargetSpotCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarge,Custo
{InstanceType=m6g.xlarge, CustomAmiId=ami-345678}']
```

Administración de actualizaciones de repositorio de paquetes de la AMI

En el primer arranque, Amazon Linux se AMIs conecta de forma predeterminada a los repositorios de paquetes para instalar las actualizaciones de seguridad antes de que se inicien otros servicios. En función de sus requisitos, puede optar por desactivar estas actualizaciones al especificar una AMI personalizada para Amazon EMR. La opción de deshabilitar esta característica solo está disponible

cuando se utiliza una AMI personalizada. De forma predeterminada, las actualizaciones del kernel de Amazon Linux y otros paquetes de software que requieren un reinicio no se actualizan. Tenga en cuenta que la configuración de red debe permitir la salida de HTTP y HTTPS a los repositorios de Amazon Linux en Amazon S3; de lo contrario, las actualizaciones de seguridad no se realizarán correctamente.

> **Warning**
>
> Recomendamos encarecidamente que elija actualizar todos los paquetes instalados durante el reinicio al especificar una AMI personalizada. Si se elige no actualizar paquetes se crean riesgos de seguridad adicionales.

Con el AWS Management Console, puede seleccionar la opción para deshabilitar las actualizaciones al elegir una AMI personalizada.

Con el AWS CLI, puede especificar `--repo-upgrade-on-boot NONE` y `--custom-ami-id` cuándo utilizar el create-cluster comando.

Con la API de Amazon EMR, puede especificar el `NONE` RepoUpgradeOnBootparámetro.

Creación de una AMI personalizada con un volumen de dispositivo raíz de Amazon EBS cifrado

Para cifrar el volumen de dispositivo raíz de Amazon EBS de una AMI de Amazon Linux para Amazon EMR, copie una imagen de instantánea desde una AMI sin cifrar a un destino cifrado. Para obtener información sobre la creación de volúmenes de EBS cifrados, consulte el cifrado de Amazon EBS en la Guía EC2 del usuario de Amazon. La AMI de origen para la instantánea puede ser la AMI base de Amazon Linux o puede copiar una instantánea de una AMI derivada de la AMI base de Amazon Linux que haya personalizado.

> **Note**
>
> A partir de la versión 5.24.0 de Amazon EMR, puede utilizar una opción de configuración de seguridad para cifrar los volúmenes de almacenamiento y los dispositivos raíz de EBS si lo especifica como proveedor de claves. AWS KMS Para obtener más información, consulte Cifrado de disco local.

Puede utilizar un proveedor de claves externo o una clave AWS KMS para cifrar el volumen raíz de EBS. Se debe permitir al rol de servicio que utiliza Amazon EMR (normalmente, el `EMR_DefaultRole` predeterminado) que cifre y descifre el volumen, como mínimo, para que Amazon EMR pueda crear un clúster con la AMI. Si se utiliza AWS KMS como proveedor de claves, esto significa que se deben permitir las siguientes acciones:

- `kms:encrypt`
- `kms:decrypt`
- `kms:ReEncrypt*`
- `kms:CreateGrant`
- `kms:GenerateDataKeyWithoutPlaintext"`
- `kms:DescribeKey"`

La forma más sencilla de hacerlo consiste en agregar el rol como usuario clave, tal y como se describe en el siguiente tutorial. Se proporciona la siguiente instrucción de política de ejemplo por si tiene que personalizar políticas de rol.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "EmrDiskEncryptionPolicy",
      "Effect": "Allow",
      "Action": [
        "kms:Encrypt",
        "kms:Decrypt",
        "kms:ReEncrypt*",
        "kms:CreateGrant",
        "kms:GenerateDataKeyWithoutPlaintext",
        "kms:DescribeKey"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
```

```
}
```

Tutorial: Creación de una AMI personalizada con un volumen de dispositivo raíz cifrado utilizando una clave de KMS

El primer paso de este ejemplo consiste en encontrar el ARN de una clave de KMS o crear una nueva. Para más información sobre la creación de claves, consulte Creación de claves en la Guía para desarrolladores de AWS Key Management Service . El siguiente procedimiento muestra cómo añadir el rol de servicio predeterminado, `EMR_DefaultRole`, como un usuario clave a la política de claves. Anote el valor ARN para la clave en el momento de crearla o editarla. Utilice el ARN más superior al crear la AMI.

Para añadir el rol de servicio de Amazon EC2 a la lista de usuarios de claves de cifrado con la consola

1. Inicia sesión en la consola AWS Key Management Service (AWS KMS) AWS Management Console y ábrela en https://console.aws.amazon.com/kms.
2. Para cambiarla Región de AWS, usa el selector de regiones en la esquina superior derecha de la página.
3. Elija el alias de la clave de KMS que vaya a utilizar.
4. En la página de detalles de la clave, en Key Users (Usuarios de claves), seleccione Add (Añadir).
5. En el cuadro de diálogo Asociar, elija el rol de servicio de Amazon EMR. El nombre del rol predeterminado es `EMR_DefaultRole`.
6. Elija Adjuntar.

Para crear una AMI cifrada con AWS CLI

- Utilice el `aws ec2 copy-image` comando de AWS CLI para crear una AMI con un volumen de dispositivo raíz de EBS cifrado y la clave que modificó. Sustituya el valor `--kms-key-id` especificado con todo el ARN de la clave que ha creado o modificado antes.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws ec2 copy-image --source-image-id MyAmiId \
--source-region us-west-2 --name MyEncryptedEMRAmi \
--encrypted --kms-key-id arn:aws:kms:us-west-2:12345678910:key/xxxxxxxx-xxxx-xxxx-
xxxx-xxxxxxxxxxxx
```

La salida del comando proporciona el ID de la AMI que ha creado, que puede especificar al crear un clúster. Para obtener más información, consulte Utilice una sola AMI personalizada en un clúster de EMR. También puede elegir personalizar esta AMI instalando software y realizando otras configuraciones. Para obtener más información, consulte Creación de una AMI de Amazon Linux personalizada a partir de una instancia preconfigurada.

Prácticas recomendadas y consideraciones

Cuando cree una AMI personalizada para Amazon EMR, tenga en cuenta lo siguiente:

- La serie 7.x de Amazon EMR se basa en Amazon Linux 2023. Para estas versiones de Amazon EMR, debe utilizar imágenes basadas en Amazon Linux 2023 para personalizarlas. AMIs Para buscar una AMI personalizada base, consulte Búsqueda de una AMI de Linux.
- Para las versiones de Amazon EMR anteriores a 7.x, no AMIs se admite Amazon Linux 2023.
- La versión 5.30.0 y posteriores de Amazon EMR y la serie 6.x de Amazon EMR se basan en Amazon Linux 2. Para estas versiones de Amazon EMR, debe utilizar imágenes basadas en Amazon Linux 2 para personalizarlas. AMIs Para buscar una AMI personalizada base, consulte Búsqueda de una AMI de Linux.
- Para las versiones de Amazon EMR anteriores a 5.30.0 y 6.x, no se admite Amazon Linux 2. AMIs
- Debe utilizar una AMI de Amazon Linux de 64 bits. No se admite una AMI de 32 bits.
- No se admite Amazon Linux AMIs con varios volúmenes de Amazon EBS.

- Base la personalización en la AMI de Amazon Linux respaldada por EBS más reciente. Para obtener una lista de Amazon Linux AMIs y la AMI correspondiente IDs, consulte la AMI de Amazon Linux.
- No copie una instantánea de una instancia de Amazon EMR existente para crear una AMI personalizada. Esto provoca errores.
- Solo son compatibles el tipo de virtualización HVM y las instancias compatibles con Amazon EMR. Asegúrese de seleccionar la imagen de HVM y un tipo de instancia compatible con Amazon EMR a medida que realice el proceso de personalización de la AMI. Para conocer las instancias y tipos de virtualización compatibles, consulte Tipos de instancias admitidas con Amazon EMR.
- El rol de servicio deben tener permisos de lanzamiento en la AMI, por lo que la AMI debe ser pública o debe ser el propietario de la AMI o que el propietario la haya compartido con usted.
- La creación de usuarios en la AMI con el mismo nombre que las aplicaciones provoca errores (por ejemplo, `hadoop`, `hdfs`, `yarn` o `spark`).
- El contenido de `/tmp`, `/var` y `/emr` (si se encuentran en la AMI), se traslada a `/mnt/tmp`, `/mnt/var` y `/mnt/emr` respectivamente durante el inicio. Los archivos se conservan, pero si hay una gran cantidad de datos, el startup puede tardar más de lo esperado.
- Si utiliza una AMI personalizada de Amazon Linux basada en una AMI de Amazon Linux con una fecha de creación del 11 de agosto de 2018, el servidor de Oozie no podrá iniciarse. Si utiliza Oozie, cree una AMI personalizada basada en un ID de AMI de Amazon Linux con una fecha de creación diferente. Puede utilizar el siguiente AWS CLI comando para obtener una lista de imágenes IDs para todas las HVM de Amazon Linux AMIs con una versión 2018.03, junto con la fecha de lanzamiento, de forma que pueda elegir una AMI de Amazon Linux adecuada como base. MyRegion Sustitúyalo por tu identificador de región, como us-west-2.

```
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
```

- En los casos en los que utilice una VPC con un nombre de dominio y un AmazonProvided DNS no estándar, no debe utilizar la `rotate` opción en la configuración de DNS del sistema operativo.
- Si crea una AMI personalizada que incluya el agente Amazon EC2 Systems Manager (SSM), el agente SSM activado puede provocar un error de aprovisionamiento en el clúster. Para evitarlo, desactive el agente SSM cuando utilice una AMI personalizada. Para ello, cuando elija y lance su EC2 instancia de Amazon, desactive el agente SSM antes de usar la instancia para crear una AMI personalizada y, posteriormente, crear su clúster de EMR.

Para obtener más información, consulte Creación de una AMI de Linux respaldada por Amazon EBS en la Guía EC2 del usuario de Amazon.

Cambio de versión de Amazon Linux al crear un clúster de EMR

Cuando lanza un clúster con Amazon EMR 6.6.0 o una versión posterior, se utiliza automáticamente la versión más reciente de Amazon Linux 2 que se haya validado para la AMI de Amazon EMR predeterminada. Puede especificar una versión de Amazon Linux diferente para su clúster con la consola Amazon EMR o AWS CLI.

Amazon EMR console

Para cambiar la versión de Amazon Linux al crear un clúster desde la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Versión de EMR, elija emr-6.6.0 o una superior.
4. En Opciones del sistema operativo, selecciona la versión de Amazon Linux, desmarca la casilla Aplicar automáticamente las últimas actualizaciones de Amazon Linux y selecciona la versión de Amazon Linux que desees.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para cambiar la versión de Amazon Linux al crear un clúster con la AWS CLI

- Utilice el parámetro `--os-release-label` para especificar la versión de Amazon Linux al ejecutar el comando `aws emr` create-cluster.

```
aws emr create-cluster --name "Cluster with Different Amazon Linux Release" \
--os-release-label 2.0.20210312.1 \
--release-label emr-6.6.0 --use-default-roles \
--instance-count 2 --instance-type m5.xlarge
```

Personalización del volumen de dispositivo raíz de Amazon EBS

Puede configurar el tipo de volumen y otros atributos, según su caso de uso y sus requisitos de costes. Puede aceptar los valores predeterminados o realizar personalizaciones.

Valores predeterminados del volumen raíz de EBS

A partir de la versión 4.x de Amazon EMR, puede especificar el tamaño del volumen raíz al crear un clúster. A partir de la versión 6.15.0 de Amazon EMR, también puede especificar las IOPS y el rendimiento del volumen raíz. Los atributos se aplican solo al volumen del dispositivo raíz de Amazon EBS y se aplican a todas las instancias del clúster. No se aplican a volúmenes de almacenamiento, que se especifican de forma independiente para cada tipo de instancia al crear el clúster.

- El tamaño del volumen raíz predeterminado es de 15 GiB a partir de la versión 6.10.0 de Amazon EMR. Las versiones anteriores tienen un tamaño de volumen raíz predeterminado de 10 GiB. Puede ajustarlo hasta 100 GiB.
- El valor predeterminado de IOPS del volumen raíz es de 3000. Puede ajustarlo hasta 16 000.
- El rendimiento predeterminado del volumen raíz es de 125. MiB/s. You can adjust this up to 1000 Mib/s

> **Note**
>
> El tamaño del volumen raíz y las IOPS no pueden tener una relación superior a 1 volumen por 500 IOPS (1:500), mientras que las IOPS y el rendimiento del volumen raíz no pueden tener una relación superior a 1 IOPS por 0,25 de rendimiento (1:0,25).

Para obtener más información sobre Amazon EBS, consulte Amazon EC2 Root Device Volume.

Tipo de volumen del dispositivo raíz con la AMI predeterminada

Cuando utiliza la AMI predeterminada, el tipo de volumen de dispositivo raíz lo determina la versión de Amazon EMR que utilice.

- Con las versiones 6.15.0 y posteriores de Amazon EMR, Amazon EMR adjunta SSD de uso general (gp3) como el tipo de volumen de dispositivo raíz.
- Con versiones de Amazon EMR anteriores a la 6.15.0, Amazon EMR adjunta el volumen SSD de uso general (gp2) como el tipo de volumen de dispositivo raíz.

Tipo de volumen del dispositivo raíz con la AMI personalizada

Una AMI personalizada puede tener otros tipos distintos de volúmenes de dispositivo raíz. Amazon EMR siempre utiliza el tipo de volumen de la AMI personalizada.

- Con las versiones 6.15.0 y posteriores de Amazon EMR, puede configurar el tamaño del volumen raíz, las IOPS y el rendimiento de su AMI personalizada, siempre que estos atributos se apliquen al tipo de volumen de la AMI personalizada.
- Con las versiones de Amazon EMR anteriores a la 6.15.0, solo puede configurar el tamaño del volumen raíz para su AMI personalizada.

Si no configura el tamaño del volumen raíz, las IOPS o el rendimiento al crear el clúster, Amazon EMR utilizará los valores de la AMI personalizada, si procede. Si decide configurar estos valores al crear el clúster, Amazon EMR utilizará los valores que especifique siempre que sean compatibles con el volumen raíz de la AMI personalizada. Para obtener más información, consulte Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR.

Precios del tamaño del volumen de dispositivo raíz

El costo del volumen de dispositivo raíz de EBS se prorratea por hora, en función de los cargos mensuales de EBS para ese tipo de volumen en la región en la que se ejecuta el clúster. Lo mismo ocurre con los volúmenes de almacenamiento. Los cargos se muestran en GB, pero debe especificar el tamaño del volumen raíz en GiB, por lo que es posible que desee tener en cuenta esto en las estimaciones (1 GB es 0,931323 GiB).

Los volúmenes SSD de uso general gp2 y gp3 se facturan de forma diferente. Para estimar los cargos asociados con los volúmenes de dispositivo raíz de EBS en su clúster, utilice las siguientes fórmulas:

SSD de uso general gp2

El precio de gp2 incluye solo el tamaño del volumen de EBS en GB.

```
($EBS size in GB/month) * 0.931323 / 30 / 24 * EMR_EBSRootVolumesizeInGiB *
 InstanceCount
```

Por ejemplo, tome un clúster que tenga un nodo principal, un nodo central y que utilice la AMI base de Amazon Linux con el volumen de dispositivo raíz de 10 GiB predeterminado. Si el costo

de EBS en la región es de 0,10 USD (GB/month, that works out to be approximately $0.00129 per instance per hour, and $0.00258 per hour for the cluster ($0.10/GB/monthdividido por 30 días, dividido por 24 horas, multiplicado por 10 GB, multiplicado por 2 instancias de clúster).

SSD de uso general gp3

El costo de la gp3 incluye el tamaño del volumen de EBS en GB, las IOPS superiores a 3000 (3000 IOPS gratuitas) y el rendimiento superior a 125 (gratuitas). MB/s (125 MB/s

```
($EBS size in GB/month) * 0.931323 / 30 / 24 * EMR_EBSRootVolumesizeInGiB *
 InstanceCount
+
($EBS IOPS/Month)/30/24* (EMR_EBSRootVolumeIops - 3000) * InstanceCount
+
($EBS throughput/Month)/30/24* (EMR_EBSRootVolumeThroughputInMb/s - 125) *
 InstanceCount
```

Por ejemplo, tome un clúster que tenga un nodo principal, un nodo central y que utilice la AMI base de Amazon Linux con el tamaño del volumen de dispositivo raíz de 15 GiB predeterminado, 4000 IOPS y 140 de rendimiento. Si el coste de EBS en la región es de 0,10 USD o más de 125. GB/month, $0.005/provisioned IOPS/month over 3000, and $0.040/provisioned MB/s/month esto equivale aproximadamente a 0,009293 USD por instancia por hora y 0,018586 USD por hora para el clúster.

Especificación de ajustes del volumen de dispositivo raíz personalizados

> **Note**
>
> El tamaño del volumen raíz y las IOPS no pueden tener una relación superior a 1 volumen por 500 IOPS (1:500), mientras que las IOPS y el rendimiento del volumen raíz no pueden tener una relación superior a 1 IOPS por 0,25 de rendimiento (1:0,25).

Console

Para especificar los atributos del volumen de dispositivo raíz de Amazon EBS desde la consola de Amazon EMR

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.

2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. Seleccione la versión 6.15.0 Amazon EMR o una posterior.
4. En Configuración del clúster, vaya a la sección Volumen raíz de EBS e introduzca un valor para cualquiera de los atributos que desee configurar.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

CLI

Para especificar los atributos del volumen de dispositivo raíz de Amazon EBS con la AWS CLI

- Utilice los parámetros `--ebs-root-volume-size`, `--ebs-root-volume-iops` y `--ebs-root-volume-throughput` del comando create-cluster, tal y como se muestra en el siguiente ejemplo.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --release-label emr-6.15.0\
--ebs-root-volume-size 20 \
--ebs-root-volume-iops 3000\
--ebs-root-volume-throughput 135\
--instance-groups InstanceGroupType=MASTER,\
InstanceCount=1,InstanceType=m5.xlarge
 InstanceGroupType=CORE,InstanceCount=2,InstanceType=m5.xlarge
```

Configure las aplicaciones al lanzar su clúster de Amazon EMR

Cuando selecciona una versión de software, Amazon EMR utiliza una imagen de máquina de Amazon (AMI) con Amazon Linux para instalar el software que usted elija al lanzar un clúster, como Hadoop, Spark y Hive. Amazon EMR ofrece nuevas versiones de forma periódica, agregando nuevas

características y aplicaciones, además de actualizaciones generales. Le recomendamos que utilice la última versión para lanzar el clúster siempre que sea posible. La última versión es la opción predeterminada al lanzar un clúster desde la consola.

Para más información sobre las versiones de Amazon EMR y las versiones de software disponibles en cada lanzamiento, consulte la Guía de versiones de Amazon EMR. Para más información sobre cómo editar las configuraciones predeterminadas de aplicaciones y software instalados en su clúster, consulte Configuración de aplicaciones en la Guía de versiones de Amazon EMR. Algunas versiones de los componentes del ecosistema de Hadoop y Spark de código abierto que se incluyen en las versiones de Amazon EMR tienen parches y mejoras, que se documentan en la Guía de versiones de Amazon EMR.

Además del software y las aplicaciones estándar que están disponibles para su instalación en su clúster, puede utilizar acciones de arranque para instalar software personalizado. Las acciones de arranque son scripts que se ejecutan en las instancias cuando se lanza el clúster y que se ejecutan en nuevos nodos que se añaden a su clúster cuando se crean. Las acciones de Bootstrap también son útiles para invocar AWS CLI comandos en cada nodo para copiar objetos de Amazon S3 a cada nodo del clúster.

> **Note**
>
> Las acciones de arranque se utilizan de forma distinta en Amazon EMR versión 4.x y posteriores. Para más información sobre estas diferencias respecto a las versiones 2.x y 3.x de la AMI de Amazon EMR, consulte Diferencias introducidas en la versión 4.x en la Guía de versiones de Amazon EMR.

Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR

Puede utilizar una acción de arranque para instalar software adicional o personalizar la configuración de las instancias de clúster. Las acciones de arranque son scripts que se ejecutan en el clúster después de que Amazon EMR lanza la instancia mediante la imagen de máquina de Amazon (AMI) de Amazon Linux. Las acciones de arranque se ejecutan antes de que Amazon EMR instale las aplicaciones que se especifican al crear el clúster y antes de que los nodos del clúster comiencen a procesar los datos. Si añade nodos a un clúster en ejecución, las acciones de arranque también se ejecutan en esos nodos de la misma manera. Puede crear acciones de arranque personalizadas y especificarlas al crear el clúster.

La mayoría de las acciones de arranque predefinidas para las versiones 2.x y 3.x de la AMI de Amazon EMR no se admiten en sus versiones 4.x. Por ejemplo, `configure-Hadoop` y `configure-daemons` no se admiten en la versión 4.x de Amazon EMR. En su lugar, Amazon EMR versión 4.x proporciona esta funcionalidad de forma nativa. Para más información sobre cómo migrar las acciones de arranque de las versiones 2.x y 3.x de la AMI de Amazon EMR a la versión 4.x de Amazon EMR, consulte Personalización de la configuración de clústeres y aplicaciones con versiones anteriores de la AMI de Amazon EMR en la Guía de versiones de Amazon EMR.

Conceptos básicos de las acciones de arranque

Las acciones de arranque se ejecutan como usuario de Hadoop de forma predeterminada. Puede ejecutar una acción de arranque con privilegios raíz utilizando `sudo`.

Todas las interfaces de administración de Amazon EMR admiten acciones de arranque. Puede especificar hasta 16 acciones de arranque por clúster proporcionando varios `bootstrap-actions` parámetros desde la consola o la API AWS CLI.

Desde la consola de Amazon EMR, puede especificar opcionalmente una acción de arranque al crear un clúster.

Cuando utilice la CLI, puede pasar referencias a scripts de acciones de arranque a Amazon EMR si agrega el parámetro `--bootstrap-actions` al crear el clúster mediante el comando `create-cluster`.

```
--bootstrap-actions Path="s3://amzn-s3-demo-bucket/filename",Args=[arg1,arg2]
```

Si la acción de arranque devuelve un código de error distinto de cero, Amazon EMR lo trata como un error y termina la instancia. Si hay demasiadas instancias que generen error en sus acciones de arranque, entonces Amazon EMR termina el clúster. Si solo unas pocas instancias generan error, Amazon EMR intenta reasignar las instancias erróneas y continuar. Utilice el código de error `lastStateChangeReason` del clúster para identificar los errores provocados por una acción de arranque.

Ejecución condicional de una acción de arranque

Para ejecutar solo una acción de arranque en el nodo maestro, puede usar una acción de arranque personalizada con un poco de lógica para determinar si el nodo es maestro.

```
#!/bin/bash
```

```
if grep isMaster /mnt/var/lib/info/instance.json | grep false;
then
    echo "This is not master node, do nothing,exiting"
    exit 0
fi
echo "This is master, continuing to execute script"
# continue with code logic for master node below
```

El siguiente resultado se imprimirá desde un nodo principal.

```
This is not master node, do nothing, exiting
```

El siguiente resultado se imprimirá desde el nodo maestro.

```
This is master, continuing to execute script
```

Para utilizar esta lógica, cargue su acción de arranque, incluido el código anterior, en su bucket de Amazon S3. En AWS CLI, añada el `--bootstrap-actions` parámetro a la llamada a la `aws emr create-cluster` API y especifique la ubicación del script de arranque como el valor de. `Path`

Acciones de apagado

Una script de acción de arranque puede crear una o más acciones de apagado escribiendo scripts en el directorio `/mnt/var/lib/instance-controller/public/shutdown-actions/`. Cuando un clúster se termina, todos los scripts en este directorio se ejecutan en paralelo. Cada script se debe ejecutar y completar en un plazo de 60 segundos.

No se garantiza la ejecución de los scripts de acción de apagado si el nodo termina con un error.

> **Note**
>
> Cuando se utiliza la versión 4.0 y posteriores de Amazon EMR, debe crear manualmente el directorio `/mnt/var/lib/instance-controller/public/shutdown-actions/` en el nodo maestro. No existe de forma predeterminada; sin embargo, después de crearlos, los scripts en este directorio se ejecutan a pesar de todo antes del apagado. Para obtener más información acerca de la conexión al nodo principal para crear directorios, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.

Usar acciones de arranque personalizadas

Puede crear un script personalizado para realizar una acción de arranque personalizada. Cualquiera de las interfaces de Amazon EMR puede hacer referencia a una acción de arranque personalizada.

> **Note**
>
> Para obtener el mejor rendimiento, le recomendamos que almacene las acciones de arranque personalizadas, los scripts y otros archivos que desee utilizar con Amazon EMR en un bucket de Amazon S3 que se encuentre en el Región de AWS mismo lugar que su clúster.

Contenido

- Agregar acciones de arranque personalizadas
- Usar una acción de arranque personalizada para copiar un objeto de Amazon S3 en cada nodo

Agregar acciones de arranque personalizadas

Console

Para crear un clúster con una acción de arranque mediante la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Acciones de arranque, elija Agregar para especificar un nombre, una ubicación del script y argumentos opcionales para la acción. Seleccione Agregar acción de arranque.
4. Opcionalmente, agregue más acciones de arranque.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

CLI

Para crear un clúster con una acción de arranque personalizada con el AWS CLI

Cuando utilice la AWS CLI para incluir una acción de arranque, especifique la `Path` y `Args` como una lista separada por comas. El siguiente ejemplo no utiliza una lista de argumentos.

- Para lanzar un clúster con una acción de arranque personalizada, escriba el siguiente comando y *`myKey`* sustitúyalo por el nombre del par de EC2 claves. Incluya `--bootstrap-actions` como parámetro y especifique la ubicación del script de arranque como valor de `Path`.

  - Usuarios de Linux, UNIX y Mac OS X:

    ```
    aws emr create-cluster --name "Test cluster" --release-label emr-4.0.0 \
    --use-default-roles --ec2-attributes KeyName=myKey \
    --applications Name=Hive Name=Pig \
    --instance-count 3 --instance-type m5.xlarge \
    --bootstrap-actions Path="s3://elasticmapreduce/bootstrap-actions/download.sh"
    ```

  - Usuarios de Windows:

    ```
    aws emr create-cluster --name "Test cluster" --release-label emr-4.2.0 --use-
    default-roles --ec2-attributes KeyName=myKey --applications Name=Hive Name=Pig
     --instance-count 3 --instance-type m5.xlarge --bootstrap-actions Path="s3://
    elasticmapreduce/bootstrap-actions/download.sh"
    ```

  Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos básicos. Todos los nodos utilizarán el tipo de instancia especificado en el comando.

  > **Note**
  >
  > Si no ha creado anteriormente el rol de servicio Amazon EMR y el perfil de EC2 instancia predeterminados, escriba `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Usar una acción de arranque personalizada para copiar un objeto de Amazon S3 en cada nodo

Puede utilizar una acción de arranque para copiar objetos desde Amazon S3 en cada nodo de un clúster antes de que se instalen las aplicaciones. AWS CLI Se instala en cada nodo de un clúster, por lo que la acción de bootstrap puede invocar AWS CLI comandos.

En el siguiente ejemplo, se muestra un script sencillo de acción de arranque que copia el archivo `myfile.jar` desde Amazon S3 en la carpeta local `/mnt1/myfolder` de cada nodo del clúster. El script se guarda en Amazon S3 con el nombre de archivo `copymyfile.sh` y contiene lo siguiente.

```
#!/bin/bash
aws s3 cp s3://amzn-s3-demo-bucket/myfilefolder/myfile.jar /mnt1/myfolder
```

Al lanzar el clúster, debe especificar el script. El siguiente AWS CLI ejemplo lo demuestra:

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.10.0 \
--use-default-roles --ec2-attributes KeyName=myKey \
--applications Name=Hive Name=Pig \
--instance-count 3 --instance-type m5.xlarge \
--bootstrap-actions Path="s3://amzn-s3-demo-bucket/myscriptfolder/copymyfile.sh"
```

Configuración del hardware y las redes de los clústeres de Amazon EMR

Una consideración importante a la hora de crear un clúster de Amazon EMR es la forma en que se configuran las EC2 instancias de Amazon y las opciones de red. En este capítulo, se abordan las siguientes opciones y, a continuación, se combinan todas ellas con las directrices y prácticas recomendadas.

- Tipos de nodos: EC2 las instancias de Amazon de un clúster de EMR se organizan en tipos de nodos. Existen tres tipos: el nodo principal, el nodo básico y los nodos de tarea. Cada tipo de nodo realiza un conjunto de roles definidos por las aplicaciones distribuidas que se instalan en el clúster. Durante un trabajo de Hadoop MapReduce o Spark, por ejemplo, los componentes de los nodos

principales y de tareas procesan los datos, transfieren la salida a Amazon S3 o HDFS y devuelven los metadatos de estado al nodo principal. Para un clúster de un solo nodo, todos los componentes se ejecutan en el nodo principal. Para obtener más información, consulte Descripción de los tipos de nodos en Amazon EMR: principales, básicos y de tarea.

- EC2 instancias: cuando creas un clúster, eliges las EC2 instancias de Amazon en las que se ejecutará cada tipo de nodo. El tipo de EC2 instancia determina el perfil de procesamiento y almacenamiento del nodo. La elección de la EC2 instancia de Amazon para sus nodos es importante porque determina el perfil de rendimiento de los tipos de nodos individuales de su clúster. Para obtener más información, consulte Configurar los tipos de EC2 instancias de Amazon para utilizarlos con Amazon EMR.
- Redes: puede lanzar su clúster de Amazon EMR en una VPC mediante una subred pública, una subred privada o una subred compartida. Su configuración de red determina cómo los clientes y los servicios pueden conectarse a los clústeres para realizar su trabajo, cómo se conectan los clústeres a los almacenes de datos y otros recursos de AWS , y las opciones de las que dispone para controlar el tráfico en esas conexiones. Para obtener más información, consulte Configuración de redes en una VPC para Amazon EMR.
- Agrupación de instancias: el conjunto de EC2 instancias que aloja cada tipo de nodo se denomina flota de instancias o grupo de instancias uniforme. La configuración de agrupación de instancias es una elección que se toma al crear un clúster. Esta opción determina cómo puede agregar nodos al clúster mientras está en ejecución. La configuración se aplica a todos los tipos de nodos. No se puede cambiar más adelante. Para obtener más información, consulte Creación de un clúster de Amazon EMR con flotas de instancias o grupos de instancias uniformes.

> **Note**
>
> La configuración de las flotas de instancias está disponible solo en las versiones 4.8.0 y posteriores de Amazon EMR, excluidas las versiones 5.0.0 y 5.0.3.

Descripción de los tipos de nodos en Amazon EMR: principales, básicos y de tarea

Utilice esta sección para entender cómo Amazon EMR utiliza cada uno de estos tipos de nodos y como base para la planificación de capacidad de los clústeres.

Nodo principal

El nodo principal administra el clúster y normalmente ejecuta los componentes principales de las aplicaciones distribuidas. Por ejemplo, el nodo principal ejecuta el ResourceManager servicio YARN para administrar los recursos de las aplicaciones. También ejecuta el NameNode servicio HDFS, realiza un seguimiento del estado de los trabajos enviados al clúster y supervisa el estado de los grupos de instancias.

Para supervisar el progreso de un clúster e interactuar directamente con las aplicaciones, puede conectar con el nodo principal a través de SSH como usuario de Hadoop. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH. La conexión con el nodo principal le permite acceder directamente a los directorios y los archivos, como los archivos de registro de Hadoop. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR. También puede ver las interfaces de usuario que publican las aplicaciones como sitios web que se ejecutan en el nodo principal. Para obtener más información, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

> **Note**
>
> Con Amazon EMR 5.23.0 y versiones posteriores, puede lanzar un clúster con tres nodos principales para admitir una alta disponibilidad de aplicaciones como YARN Resource Manager, HDFS, Spark NameNode, Hive y Ganglia. El nodo principal ya no es un posible punto único de error con esta característica. Si se produce un error en uno de los nodos principales, se produce un error automáticamente en Amazon EMR a través de un nodo principal en espera y se sustituye el nodo principal con error por uno nuevo con las mismas acciones de arranque y configuración. Para más información, consulte Planificación y configuración de nodos principales.

Nodos básicos

Los nodos básicos los administra el nodo principal. Los nodos secundarios ejecutan el daemon del nodo de datos para coordinar el almacenamiento de datos como parte del Sistema de archivos distribuido de Hadoop (HDFS). También ejecutan el daemon Task Tracker y realizan otras tareas de cálculo en paralelo en los datos que las aplicaciones instaladas requieren. Por ejemplo, un nodo principal ejecuta los NodeManager daemons de YARN, las tareas de Hadoop y los ejecutores de Spark. MapReduce

Solo hay un grupo de instancias principal o una flota de instancias por clúster, pero puede haber varios nodos ejecutándose en varias instancias de Amazon en el grupo de EC2 instancias o en la flota de instancias. Con los grupos de instancias, puedes añadir y eliminar EC2 instancias de Amazon mientras el clúster está en ejecución. También puede configurar el escalado automático para agregar instancias en función del valor de una métrica. Para obtener más información sobre cómo añadir y eliminar EC2 instancias de Amazon con la configuración de grupos de instancias, consultaUtilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes.

Con las flotas de instancias, puede agregar y eliminar instancias de forma eficaz modificando las capacidades de destino de la flota de instancias para las opciones bajo demanda y de spot según corresponda. Para obtener más información acerca de las capacidades de destino, consulte Opciones de flota de instancias.

> **Warning**
>
> La eliminación de daemons de HDFS de un nodo secundario de ejecución o la terminación de nodos secundarios conlleva el riesgo de pérdida de datos. Tenga cuidado al configurar nodos secundarios que utilizan instancias de spot. Para obtener más información, consulte ¿Cuándo se deben utilizar las instancias de spot?.

Nodos de tarea

Puede utilizar los nodos de tareas para añadir potencia y realizar tareas de cálculo en paralelo con los datos, como las tareas de Hadoop y los ejecutores de MapReduce Spark. Los nodos de tareas no ejecutan el daemon de nodo de datos, ni tampoco almacenan datos en HDFS. Al igual que con los nodos principales, puede añadir nodos de tareas a un clúster añadiendo EC2 instancias de Amazon a un grupo de instancias uniforme existente o modificando las capacidades objetivo de una flota de instancias de tareas.

Con la configuración de grupo de instancias uniforme, puede tener hasta un total de 48 grupos de instancias de tarea. La posibilidad de añadir grupos de instancias de esta manera te permite combinar tipos de EC2 instancias de Amazon y opciones de precios, como instancias bajo demanda e instancias puntuales. Esto le ofrece flexibilidad para responder a los requisitos de carga de trabajo de forma rentable.

Con la configuración de flota de instancias, la capacidad de combinar tipos de instancia y opciones de compra está integrada, por lo que solo hay una flota de instancias de tarea.

Dado que las instancias de spot se utilizan a menudo para ejecutar nodos de tarea, Amazon EMR tiene una funcionalidad predeterminada para programar trabajos de YARN, de modo que los trabajos en ejecución no presenten errores cuando los nodos de tarea que se ejecutan en las instancias de spot se terminen. Para ello, Amazon EMR permite que los procesos maestros de la aplicación se ejecuten únicamente en los nodos principales. El proceso maestro de la aplicación controla los trabajos en ejecución y debe mantenerse activo durante toda la vida del trabajo.

La versión 5.19.0 y posteriores de Amazon EMR utilizan la característica integrada de etiquetas de nodo YARN para lograrlo. (Las versiones anteriores utilizaban una revisión de código). Las propiedades en las clasificaciones de configuración `yarn-site` y `capacity-scheduler` se ajustan de forma predeterminada para que capacity-scheduler y fair-scheduler de YARN utilicen las etiquetas de nodo. Amazon EMR etiqueta automáticamente los nodos principales con la etiqueta `CORE` y establece las propiedades para que los maestros de la aplicación se programen únicamente en los nodos con la etiqueta CORE. La modificación manual de las propiedades relacionadas en las clasificaciones de configuración yarn-site y capacity-scheduler o directamente en los archivos XML asociados podría interrumpir esta característica o modificar esta funcionalidad.

A partir de la serie de versiones 6.x de Amazon EMR, la característica de etiquetas de nodo YARN está desactivada de forma predeterminada. De forma predeterminada, los procesos principales de la aplicación se pueden ejecutar tanto en nodos básicos como en nodos de tarea. Puede habilitar la función de etiquetas de nodo YARN configurando las siguientes propiedades:

- `yarn.node-labels.enabled: true`
- `yarn.node-labels.am.default-node-label-expression: 'CORE'`

A partir de la serie de versiones 7.x de Amazon EMR, Amazon EMR asigna etiquetas de nodo YARN a las instancias según su tipo de mercado, como “Bajo demanda” o “Spot”. Puede habilitar las etiquetas de nodo y restringir los procesos de aplicación a ON_DEMAND configurando las siguientes propiedades:

```
yarn.node-labels.enabled: true
yarn.node-labels.am.default-node-label-expression: 'ON_DEMAND'
```

Si utiliza Amazon EMR 7.0 o una versión posterior, puede restringir el proceso de aplicación a los nodos con la etiqueta `CORE` mediante la siguiente configuración:

```
yarn.node-labels.enabled: true
yarn.node-labels.am.default-node-label-expression: 'CORE'
```

En el caso de las versiones 7.2 y posteriores de Amazon EMR, si el clúster utiliza el escalado administrado con etiquetas de nodo, Amazon EMR intentará escalar el clúster en función del proceso de aplicación y de la demanda del ejecutor de forma independiente.

Por ejemplo, si utiliza la versión 7.2 o posteriores de Amazon EMR y restringe el proceso de aplicación a los nodos `ON_DEMAND`, el escalado administrado escala verticalmente los nodos `ON_DEMAND` si aumenta la demanda del proceso de aplicación. Del mismo modo, si restringes el proceso de solicitud a `CORE` los nodos, el escalado gestionado amplía `CORE` los nodos si aumenta la demanda del proceso de solicitud.

Para obtener información sobre propiedades específicas, consulte Configuración de Amazon EMR para evitar errores en los trabajos debido a la terminación de instancias de spot de los nodos de tarea.

Configurar los tipos de EC2 instancias de Amazon para utilizarlos con Amazon EMR

EC2 las instancias vienen en diferentes configuraciones, conocidas como tipos de instancias. Cada tipo de instancia tiene una CPU, entrada/salida y capacidad de almacenamiento diferentes. Además del tipo de instancia, puedes elegir diferentes opciones de compra para las EC2 instancias de Amazon. Puede especificar diferentes tipos de instancia y opciones de compra dentro de grupos o flotas de instancias. Para obtener más información, consulte Creación de un clúster de Amazon EMR con flotas de instancias o grupos de instancias uniformes. Para obtener recomendaciones sobre cómo elegir los tipos de instancias y las opciones de compra para su aplicación, consulte Configuración de tipos de instancias de clúster de Amazon EMR y prácticas recomendadas para instancias de Spot.

> **Important**
>
> Cuando eliges un tipo de instancia mediante el AWS Management Console, el número de vCPU que se muestra para cada tipo de instancia es el número de núcleos virtuales de YARN para ese tipo de instancia, no el número de EC2 v CPUs para ese tipo de instancia. Para obtener más información sobre el número de v CPUs para cada tipo de instancia, consulte Tipos de EC2 instancias de Amazon.

Temas

- Tipos de instancias admitidas con Amazon EMR

- Configuración de redes en una VPC para Amazon EMR
- Creación de un clúster de Amazon EMR con flotas de instancias o grupos de instancias uniformes

Tipos de instancias admitidas con Amazon EMR

En esta sección se describen los tipos de instancias que admite Amazon EMR, organizados por Región de AWS. Para obtener más información sobre los tipos de instancias, consulte la matriz de tipos de EC2 instancias de Amazon y AMI de Amazon Linux.

No todos los tipos de instancias están disponibles en todas las regiones y la disponibilidad de las instancias está sujeta a la disponibilidad y la demanda en la región y la zona de disponibilidad especificadas. La zona de disponibilidad de una instancia viene determinada por la subred que utilice para lanzar el clúster.

Consideraciones

Tenga en cuenta lo siguiente al elegir los tipos de instancias del clúster de Amazon EMR.

> **Important**
>
> Cuando eliges un tipo de instancia mediante el AWS Management Console, el número de vCPU que se muestra para cada tipo de instancia es el número de núcleos virtuales de YARN para ese tipo de instancia, no el número de EC2 v CPUs para ese tipo de instancia. Para obtener más información sobre el número de v CPUs para cada tipo de instancia, consulte Tipos de EC2 instancias de Amazon.

- Si crea un clúster con un tipo de instancia que no está disponible en la región y la zona de disponibilidad especificadas, es posible que no se pueda aprovisionar el clúster o que se bloquee durante el aprovisionamiento. Para obtener información sobre la disponibilidad de las instancias, consulte la página de precios de Amazon EMR o consulte las tablas de Tipos de instancias compatibles por Región de AWS en esta página.
- A partir de la versión de 5.13.0 de Amazon EMR, todas las instancias utilizan virtualización HVM y almacenamiento respaldado por EBS para los volúmenes raíz. Cuando se utilizan versiones de Amazon EMR anteriores a la 5.13.0, algunas instancias de generaciones anteriores utilizan virtualización PVM. Para más información, consulte Tipos de virtualización de AMI de Linux.
- Debido a la falta de soporte de hardware y a la configuración predeterminada, lo que puede provocar una infrautilización de la memoria y los núcleos, no recomendamos el uso de los tipos

de instancias `c7a`, `c7i`, `m7i`, `m7i-flex`, `r7a`, `r7i`, `r7iz`, `i4i.12xlarge` y `i4i.24xlarge` si ejecuta versiones de Amazon EMR anteriores a 5.36.1 y 6.10.0. Si ejecuta estos tipos de instancias en esas versiones, es posible que experimente un rendimiento inferior y que no obtenga los beneficios esperados de los tipos de instancias más nuevos, como `c7i` frente a `c6i`. Para una utilización y un rendimiento óptimos de los recursos con estos tipos de rendimiento, debe ejecutar la versión 5.36.1 y versiones posteriores o la 6.10.0 y versiones posteriores a fin de maximizar sus capacidades.

- Algunos tipos de instancia son compatibles con redes mejoradas. Para obtener más información, consulte Redes mejoradas en Linux.
- Los controladores NVIDIA y CUDA se instalan en los tipos de instancia GPU de forma predeterminada.

Tipos de instancias compatibles por Región de AWS

En las tablas siguientes se enumeran los tipos de EC2 instancias de Amazon compatibles con Amazon EMR, organizados por. Región de AWS En las tablas también se enumeran las primeras versiones de las series 5.x, 6.x y 7.x de Amazon EMR que admiten cada tipo de instancia.

Este de EE. UU. (Norte de Virginia): us-east-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación acelerada	f2.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	f2.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	f2.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0
	g6e.2xlarge	emr-7.5.0
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	h1.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	h1.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i7i.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7ie.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i7ie.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.3xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.18xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Este de EE. UU. (Ohio): us-east-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7a.2xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7a.4xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7a.8xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7a.48xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.8xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7a.xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7a.2xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7a.4xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7a.8xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7a.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7a.16xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7a.24xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7a.32xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7a.48xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.4xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.6.0, emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0
	g6e.2xlarge	emr-7.5.0
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7a.xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7a.2xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7a.4xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7a.8xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7a.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7a.16xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7a.24xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7a.32xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7a.48xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.16xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7iz.xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7iz.2xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7iz.4xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7iz.8xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7iz.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7iz.16xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r7iz.32xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	h1.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	h1.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.1, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i7i.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7ie.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i7ie.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.3xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.18xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Oeste de EE. UU. (Norte de California): us-west-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Oeste de EE. UU. (Oregón): us-west-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación acelerada	f2.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	f2.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	f2.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6e.2xlarge	emr-7.5.0
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7iz.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	h1.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	h1.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i7i.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7i.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	i7ie.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.3xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i7ie.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.18xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

AWS GovCloud (EE. UU.-Oeste): -1 us-gov-west

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

AWS GovCloud (EEUU-Este) - -1 us-gov-east

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Optimización de almacenamiento	i3.xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.16.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

África (Ciudad del Cabo): af-south-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.18xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Asia-Pacífico (Hong Kong): ap-east-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Asia-Pacífico (Yakarta): ap-southeast-3

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5.16xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5d.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5d.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5d.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5d.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5d.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m5d.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m6g.xlarge	emr-5.30.2, emr-6.1.1, emr-7.0.0
	m6g.2xlarge	emr-5.30.2, emr-6.1.1, emr-7.0.0
	m6g.4xlarge	emr-5.30.2, emr-6.1.1, emr-7.0.0
	m6g.8xlarge	emr-5.30.2, emr-6.1.1, emr-7.0.0
	m6g.12xlarge	emr-5.30.2, emr-6.1.1, emr-7.0.0
	m6g.16xlarge	emr-5.30.2, emr-6.1.1, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i.16xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i.48xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i-flex.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i-flex.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	m7i-flex.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5.9xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5.18xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5d.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5d.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5d.9xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5d.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5d.18xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5d.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5n.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5n.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5n.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5n.9xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c5n.18xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	c6g.xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.2xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	c6g.4xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	c6g.8xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	c6g.12xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	c6g.16xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5.16xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5d.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5d.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5d.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5d.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5d.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5d.16xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r5d.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r6g.xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.2xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	r6g.4xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	r6g.8xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	r6g.12xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	r6g.16xlarge	emr-5.31.1, emr-6.1.1, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r7i.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r7i.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r7i.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r7i.16xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r7i.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	r7i.48xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3.4xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3.8xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3.16xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3en.xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3en.2xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3en.3xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3en.6xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3en.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i3en.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.30.2, emr-6.0.1, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Asia Pacífico (Melbourne) - ap-southeast-4

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5d.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5d.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5d.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5d.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5d.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5d.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m5d.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6g.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6g.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6g.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6g.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6gd.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6gd.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6gd.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6gd.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6gd.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m6gd.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i.48xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i-flex.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i-flex.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i-flex.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i-flex.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	m7i-flex.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5.9xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5.18xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5d.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5d.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5d.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5d.9xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5d.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5d.18xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c5d.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c6g.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c6g.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c6g.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c6g.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5d.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5d.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5d.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5d.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5d.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5d.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r5d.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r6g.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r6g.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r6g.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r6g.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r6g.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r7i.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r7i.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r7i.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r7i.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r7i.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r7i.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	r7i.48xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3.4xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3.8xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3.16xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3en.xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3en.2xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.3xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3en.6xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3en.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i3en.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.36.0, emr-6.8.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Asia Pacífico (Malasia) - ap-southeast-5

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m6g.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6g.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6g.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6g.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6g.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6g.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6gd.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6gd.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6gd.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6gd.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6gd.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6gd.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6i.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6i.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6i.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6i.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6i.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6i.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6i.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6id.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6id.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6id.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6id.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6id.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6id.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m6id.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7g.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7g.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7g.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7g.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7g.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7g.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7gd.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7gd.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7gd.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7gd.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7gd.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7gd.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i.48xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i-flex.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i-flex.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i-flex.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i-flex.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	m7i-flex.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
Computación optimizada	c6g.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6g.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6g.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6g.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6g.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6g.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6gn.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6gn.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6gn.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6gn.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6gn.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6i.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6i.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6i.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6i.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6i.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6i.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6i.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6id.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6in.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6in.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6in.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6in.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6in.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6in.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c6in.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7g.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7g.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7g.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7g.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7g.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7g.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7gd.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7gd.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7gd.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7gd.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7gd.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i.48xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i-flex.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i-flex.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i-flex.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i-flex.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i-flex.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	c7i-flex.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
Computación acelerada	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
Optimizada para memoria	r6g.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6g.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6g.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6g.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6g.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6g.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6i.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6i.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6i.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6i.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6i.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6i.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6i.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r6id.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7g.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7g.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7g.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7g.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7g.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7gd.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7gd.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7gd.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7gd.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7gd.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7gd.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7i.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7i.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7i.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7i.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7i.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7i.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	r7i.48xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2idn.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2idn.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2idn.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2iedn.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
Optimización de almacenamiento	i3en.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i3en.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i3en.3xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i3en.6xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i3en.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i3en.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i4i.2xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i4i.4xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i4i.8xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i4i.12xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i4i.16xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i4i.24xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0
	i4i.32xlarge	emr-5.36.2, emr-6.11.1, emr-7.0.0

Asia-Pacífico (Bombay): ap-south-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i-flex.4xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.6.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Asia-Pacífico (Hyderabad): ap-south-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.3xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.6xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Asia-Pacífico (Osaka): ap-northeast-3

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Asia-Pacífico (Seúl): ap-northeast-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0
	g6e.2xlarge	emr-7.5.0
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Asia-Pacífico (Singapur): ap-southeast-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Asia Pacífico (Sídney): ap-southeast-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	f2.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	f2.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	f2.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Asia Pacífico (Tokio): ap-northeast-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6e.2xlarge	emr-7.5.0
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7iz.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i7ie.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.3xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.18xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i7ie.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Canadá (centro): ca-central-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.2xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.4xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.8.2, emr-5.0.2, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.16xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.2, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Oeste de Canadá (Calgary): ca-west-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5d.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5d.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5d.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5d.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5d.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m5d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6g.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6g.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6g.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6g.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6g.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6g.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6gd.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6gd.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6gd.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6gd.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6gd.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6i.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c5.9xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c5.18xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6g.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6g.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6g.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6g.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6g.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6g.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6gn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6gn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6gn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6gn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6gn.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6i.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6g.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6g.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6g.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6g.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6g.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6g.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6i.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6i.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6i.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6i.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6i.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6i.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6i.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Optimización de almacenamiento	i3en.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i3en.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.3xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i3en.6xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i3en.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i3en.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

China (Ningxia): cn-northwest-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

China (Pekín): cn-north-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Europa (Fráncfort): eu-central-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	m8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	c8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0
	g6e.2xlarge	emr-7.5.0
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	gr6.8xlarge	emr-7.2.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7iz.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8gd.xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.2xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.4xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8gd.8xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.12xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.16xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.24xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	r8gd.48xlarge	emr-5.36.2, emr-6.15.0, emr-7.1.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i7ie.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.3xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.18xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i7ie.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Europa (Zúrich): eu-central-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
Computación acelerada	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
Optimizada para memoria	r5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	d3.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	d3.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	d3.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.3xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.6xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Europa (Irlanda): eu-west-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	p2.xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p2.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p2.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7a.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7a.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7iz.32xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3en.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	h1.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	h1.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	h1.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i7ie.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.3xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.18xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Europa (Londres): eu-west-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.8xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.4xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.8.2, emr-5.0.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	f2.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	f2.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	f2.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	g3.4xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g3.8xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3.16xlarge	emr-5.18.0, emr-6.0.0, emr-7.0.0
	g3s.xlarge	emr-5.19.0, emr-6.0.0, emr-7.0.0
	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p3.2xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p3.8xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p3.16xlarge	emr-5.10.0, emr-6.0.0, emr-7.0.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	z1d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.3xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	z1d.6xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	z1d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.3, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i7ie.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.3xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.6xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.18xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i7ie.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Europa (Milán): eu-south-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.9xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.2xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.29.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Europa (España): eu-south-2

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7a.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7a.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7a.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7a.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7a.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7a.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7a.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7a.48xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i.48xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i-flex.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i-flex.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i-flex.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i-flex.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i-flex.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7i-flex.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7a.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7a.48xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i.48xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i-flex.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i-flex.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i-flex.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i-flex.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c7i-flex.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i-flex.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0
	g6e.2xlarge	emr-7.5.0
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
Optimizada para memoria	r5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r7a.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7a.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7a.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7a.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7a.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7a.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7a.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7a.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7a.48xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7i.48xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.3xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.6xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	im4gn.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	im4gn.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	im4gn.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	im4gn.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	im4gn.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Europa (París): eu-west-3

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.12xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7i.xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.4xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i.24xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.9.2, emr-5.5.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.12xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	d3.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.5.3, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	im4gn.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	im4gn.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	is4gen.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Europa (Estocolmo): eu-north-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7a.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.32xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7a.48xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i-flex.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	m8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.9xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7a.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7a.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7a.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7a.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7a.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7a.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7a.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7a.32xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7a.48xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c7i.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	g6e.xlarge	emr-7.5.0
	g6e.2xlarge	emr-7.5.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g6e.4xlarge	emr-7.5.0
	g6e.8xlarge	emr-7.5.0
	g6e.12xlarge	emr-7.5.0
	g6e.16xlarge	emr-7.5.0
	g6e.24xlarge	emr-7.5.0
	g6e.48xlarge	emr-7.5.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r5dn.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5dn.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6idn.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6idn.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7a.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7a.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7a.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7a.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7a.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7a.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7a.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7a.32xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7a.48xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7i.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	r8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.17.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Israel (Tel Aviv) - il-central-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Computación optimizada	c5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gn.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gn.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gn.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gn.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6gn.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	d3.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	d3.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	d3.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	d3.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.3xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.6xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Medio Oriente (Baréin): me-south-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5.24xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5d.8xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5d.9xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5n.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5d.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5d.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	i3.2xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.24.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Medio Oriente (EAU): me-central-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6gd.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación optimizada	c5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.9xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.18xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r5d.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6g.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r6i.32xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.4xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.8xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3.16xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.2xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3en.3xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.6xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i3en.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.36.0, emr-6.7.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

América del Sur (São Paulo): sa-east-1

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Uso general	m5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	m5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	m5zn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.3xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.6xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m5zn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6g.xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.2xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.4xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.8xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.12xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0
	m6g.16xlarge	emr-5.30.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	m6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	m6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	m7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7gd.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	m7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	m7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	m7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
Computación optimizada	c5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5a.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5a.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5a.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c5ad.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5ad.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c5d.9xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.18xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	c5n.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.9xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c5n.18xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	c6a.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6a.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6a.48xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	c6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	c6gn.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	c6id.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c6id.12xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.24xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6id.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	c6in.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c6in.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7g.xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.2xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.4xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.8xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.12xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7g.16xlarge	emr-5.36.1, emr-6.7.0, emr-7.0.0
	c7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.2xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.4xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.8xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.12xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c7i-flex.16xlarge	emr-4.2.0, emr-5.0.0, emr-6.0.0, emr-7.0.0
	c8g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	c8g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.24xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	c8g.48xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
Computación acelerada	g4dn.xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.2xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.4xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.8xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.12xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g4dn.16xlarge	emr-5.30.0, emr-6.0.0, emr-7.0.0
	g5.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	g5.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.12xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g5.48xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	g6.xlarge	emr-7.2.0
	g6.2xlarge	emr-7.2.0
	g6.4xlarge	emr-7.2.0
	g6.8xlarge	emr-7.2.0
	g6.12xlarge	emr-7.2.0
	g6.16xlarge	emr-7.2.0
	g6.24xlarge	emr-7.2.0
	g6.48xlarge	emr-7.2.0
	gr6.4xlarge	emr-7.2.0
	gr6.8xlarge	emr-7.2.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	p4d.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	p5.48xlarge	emr-6.14.0, emr-7.0.0
Optimizada para memoria	r5.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5a.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.8xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5a.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.16xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5a.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.2xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.4xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.12xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5ad.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5ad.24xlarge	emr-5.20.0, emr-6.0.0, emr-7.0.0
	r5b.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5b.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5b.24xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r5d.xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.2xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.4xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.8xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.12xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.16xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0
	r5d.24xlarge	emr-5.13.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r5n.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r5n.24xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.2xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.4xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.8xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6g.12xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6g.16xlarge	emr-5.31.0, emr-6.1.0, emr-7.0.0
	r6gd.xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.2xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.4xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.8xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.12xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6gd.16xlarge	emr-5.33.0, emr-6.3.0, emr-7.0.0
	r6i.xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.2xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.4xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.8xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.12xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r6i.16xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.24xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r6i.32xlarge	emr-5.35.0, emr-6.6.0, emr-7.0.0
	r7g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.12xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	r7i.xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.2xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.4xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	r7i.8xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.16xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	r7i.48xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	x1.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x1e.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x1e.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	x1e.32xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	x2idn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2idn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.2xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.4xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.8xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.16xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.24xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
	x2iedn.32xlarge	emr-5.36.1, emr-6.9.0, emr-7.0.0
Optimización de almacenamiento	i3.xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i3.2xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.4xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.8xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3.16xlarge	emr-5.9.0, emr-6.0.0, emr-7.0.0
	i3en.xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.2xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.3xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.6xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.12xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i3en.24xlarge	emr-5.25.0, emr-6.0.0, emr-7.0.0
	i4g.xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.2xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0

Clase de instancia	Tipo de instancia	Versión mínima admitida de Amazon EMR (5.x, 6.x, 7.x)
	i4g.4xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.8xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4g.16xlarge	emr-5.36.1, emr-6.10.0, emr-7.0.0
	i4i.xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.2xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.4xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.8xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.12xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.16xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0
	i4i.24xlarge	emr-5.0.0, emr-6.0.0, emr-7.0.0
	i4i.32xlarge	emr-5.36.1, emr-6.8.0, emr-7.0.0

instancias de generaciones anteriores

Amazon EMR admite las Instancias de generaciones anteriores para admitir las aplicaciones que están optimizadas para estas instancias y que todavía no se han actualizado. Para obtener más información acerca de estos tipos de instancias y sobre las rutas de actualización, consulte Instancias de generaciones anteriores.

Clase de instancia	Tipos de instancias
General Purpose	m1.small[1] | m1.medium[1] | m1.large[1] | m1.xlarge[1] | m3.xlarge[1] | m3.2xlarge[1] | m4.large | m4.xlarge | m4.2xlarge | m4.4xlarge | m4.10xlarge | m4.16xlarge
Compute Optimized	c1.medium[1] [2] | c1.xlarge[1] | c3.xlarge[1] | c3.2xlarge[1] | c3.4xlarge[1] | c3.8xlarge[1] | c4.large | c4.xlarge | c4.2xlarge | c4.4xlarge | c4.8xlarge
Memory Optimized	m2.xlarge[1] | m2.2xlarge[1] | m2.4xlarge[1] | r3.xlarge | r3.2xlarge | r3.4xlarge | r3.8xlarge | r4.xlarge | r4.2xlarge | r4.4xlarge | r4.8xlarge | r4.16xlarge
Storage Optimized	d2.xlarge | d2.2xlarge | d2.4xlarge | d2.8xlarge | i2.xlarge | i2.2xlarge | i2.4xlarge | i2.8xlarge

[1] Utiliza la AMI de virtualización PVM con las versiones de Amazon EMR anteriores a la 5.13.0. Para obtener más información, consulte Tipos de virtualización de una AMI de Linux.

[2] No se admite en la versión 5.15.0.

Opciones de compra de instancias en Amazon EMR

Cuando configuras un clúster, eliges una opción de compra para las EC2 instancias de Amazon. Puede elegir utilizar instancias bajo demanda, instancias de spot o ambas. Los precios varían en función del tipo de instancia y de la región. El precio de Amazon EMR se suma al precio de Amazon (el EC2 precio de los servidores subyacentes) y al precio de Amazon EBS (si se adjuntan volúmenes de Amazon EBS). Para conocer los precios actuales, consulte Precios de Amazon EMR.

La opción elegida de utilizar grupos de instancias o flotas de instancias en el clúster determina cómo puede cambiar las opciones de compra de instancias mientras se ejecuta un clúster. Si eliges grupos

de instancias uniformes, solo puedes especificar la opción de compra de un grupo de instancias al crearlo, y el tipo de instancia y la opción de compra se aplican a todas las EC2 instancias de Amazon de cada grupo de instancias. Si decide utilizar flotas de instancias, puede cambiar las opciones de compra después de haber creado cada flota de instancias y puede combinar opciones de compra para satisfacer el objetivo de capacidad que especifique. Para obtener más información acerca de estas configuraciones, consulte Creación de un clúster de Amazon EMR con flotas de instancias o grupos de instancias uniformes.

instancias bajo demanda

Con las instancias bajo demanda, paga la capacidad de cómputo por segundo. De forma opcional, puede hacer que estas instancias bajo demanda utilicen las opciones de compra de instancia reservada o instancia dedicada. Con las instancias reservadas, realiza un pago único para una instancia a fin de reservar capacidad. Las instancias dedicadas están aisladas físicamente, a nivel de hardware del host, de las instancias que pertenecen a otras AWS cuentas. Para obtener más información sobre las opciones de compra, consulta Instance Purchasing Options en la Guía del EC2 usuario de Amazon.

Uso de instancias reservadas

Para usar instancias reservadas en Amazon EMR, usa Amazon EC2 para comprar la instancia reservada y especificar los parámetros de la reserva, incluido el alcance de la reserva, ya que se aplica a una región o una zona de disponibilidad. Para obtener más información, consulte Amazon EC2 Reserved Instances y Comprar instancias reservadas en la Guía del EC2 usuario de Amazon. Cuando haya adquirido una instancia reservada, y si se dan todas las condiciones siguientes, Amazon EMR la utilizará cuando se lance un clúster:

- Una instancia bajo demanda se especifica en la configuración del clúster que coincide con la especificación de instancia reservada.
- El clúster se ha lanzado dentro del ámbito de la reserva de la instancia (la zona de disponibilidad o región).
- La capacidad de instancia reservada sigue estando disponible

Por ejemplo, supongamos que adquiere una instancia reservada `m5.xlarge` con la reserva de instancia en el ámbito de la región Este de EE. UU. A continuación lanza un clúster de Amazon EMR en el Este de EE. UU. que utiliza dos instancias `m5.xlarge`. La primera instancia se factura con la tarifa de instancia reservada y la otra se factura con la tarifa de la opción bajo demanda. La capacidad de instancias reservadas se utiliza antes de crear las instancias bajo demanda.

Uso de instancias dedicadas

Para usar instancias dedicadas, debe comprar instancias dedicadas a través de Amazon EC2 y, a continuación, crear una VPC con el atributo de arrendamiento dedicado. Dentro de Amazon EMR, puede especificar que un clúster se debe lanzar en esta VPC. Cualquier instancia bajo demanda del clúster que se ajuste a la especificación de instancia dedicada usará las instancias dedicadas disponibles cuando se lanza el clúster.

> **Note**
>
> Amazon EMR no admite la configuración del atributo `dedicated` en instancias individuales.

Spot Instances

Las instancias puntuales de Amazon EMR ofrecen la opción de adquirir capacidad de EC2 instancias de Amazon a un coste reducido en comparación con la compra bajo demanda. La desventaja de utilizar instancias de spot es que las instancias pueden terminar si la capacidad de spot deja de estar disponible para el tipo de instancia que está ejecutando. Para obtener más información sobre si el uso de instancias de spot podría ser adecuado para su aplicación, consulte ¿Cuándo se deben utilizar las instancias de spot?.

Cuando Amazon EC2 tiene capacidad no utilizada, ofrece EC2 instancias a un coste reducido, denominado precio spot. El precio fluctúa en función de la disponibilidad y la demanda y se establece por región y zona de disponibilidad. Cuando elige las instancias puntuales, especifica el precio puntual máximo que está dispuesto a pagar por cada tipo de EC2 instancia. Cuando el precio de spot de la zona de disponibilidad del clúster se encuentra por debajo del precio de spot máximo especificado para dicho tipo de instancia, se lanzan las instancias. Mientras se ejecutan las instancias, se le factura el precio de spot actual, no el precio de spot máximo.

> **Note**
>
> Las instancias de spot con una duración definida (también conocidas como bloques de spot) dejarán de estar disponibles para los nuevos clientes a partir del 1 de julio de 2021. En el caso de los clientes que hayan utilizado previamente la característica, se continuará ofreciendo soporte a las instancias de spot con una duración definida hasta el 31 de diciembre de 2022.

Para ver los precios actuales, consulta los precios de Amazon EC2 Spot Instances. Para obtener más información, consulte Spot Instances en la Guía del EC2 usuario de Amazon. Al crear y configurar un clúster, especifica las opciones de red que en última instancia determinan la zona de disponibilidad donde se lanza el clúster. Para obtener más información, consulte Configuración de redes en una VPC para Amazon EMR.

> **Tip**
>
> Puede ver el precio de spot en tiempo real en la consola al pasar el ratón sobre la ayuda contextual situada junto a la opción de compra Spot al crear un clúster utilizando las Advanced Options (Opciones avanzadas). Se muestran los precios de cada zona de disponibilidad en la región seleccionada. Los precios más bajos se muestran en filas de color verde. Debido a las fluctuaciones de los precios de spot entre las zonas de disponibilidad, la selección de la zona de disponibilidad con el menor precio inicial podría no coincidir con el precio más bajo durante la vida útil del clúster. Para obtener resultados óptimos, estudie el historial de precios de la zona de disponibilidad antes elegir. Para obtener más información, consulte el historial de precios de Spot Instance en la Guía del EC2 usuario de Amazon.

Las opciones de instancias de spot dependen de si utiliza flotas de instancias o grupos de instancias uniformes en la configuración del clúster.

Instancias de subasta en grupos de instancias uniformes

Al utilizar las instancias de spot en un grupo de instancias uniforme, todas las instancias de un grupo de instancias deben ser instancias de spot. Especifique una única subred o zona de disponibilidad para el clúster. Para cada grupo de instancias, especifique una sola instancia de spot y un precio de spot máximo. Las instancias de spot de ese tipo se lanzan si el precio de spot de la zona de disponibilidad y la región del clúster está por debajo del precio de spot máximo. Las instancias se terminan si el precio de spot está por encima de su precio de spot máximo. Solo puede establecer el precio de spot al configurar un grupo de instancias. No se puede cambiar más adelante. Para obtener más información, consulte Creación de un clúster de Amazon EMR con flotas de instancias o grupos de instancias uniformes.

Instancias de subasta en flotas de instancias

Al utilizar la configuración de flotas de instancias, las opciones adicionales le ofrecen mayor control sobre cómo lanzar y terminar instancias de spot. Fundamentalmente, las flotas de instancias utilizan un método de diferente a los grupos de instancias uniformes para lanzar instancias. La forma

de funcionamiento consiste en establecer una capacidad de destino para las instancias de spot (e instancias bajo demanda) y hasta cinco tipos de instancias. También puede especificar una capacidad ponderada para cada tipo de instancia o utilizar el vCPU (vcores de YARN) del tipo de instancia como capacidad ponderada. Esta capacidad ponderada se tiene en cuenta en la capacidad de destino cuando se provisiona una instancia de ese tipo. Amazon EMR aprovisiona instancias con ambas opciones de compra hasta que se alcanza la capacidad de destino para cada destino. Además, puede definir una serie de zonas de disponibilidad de las cuales Amazon EMR pueda elegir al lanzar instancias. También proporciona opciones de spot adicionales para cada flota, incluido un tiempo de espera de aprovisionamiento. Para obtener más información, consulte Planificación y configuración de flotas de instancias para su clúster de Amazon EMR.

Comportamiento y opciones de almacenamiento de instancias en Amazon EMR

Descripción general

El almacén de instancias y el almacenamiento de volúmenes de Amazon EBS se utilizan para los datos de HDFS, así como para los búferes, cachés, datos de pruebas y otro contenido temporal que algunas aplicaciones pueden “volcar” en el sistema de archivos local.

Amazon EBS funciona de forma diferente en Amazon EMR que en las instancias normales de Amazon EC2 . Los volúmenes de Amazon EBS asociados a clústeres de Amazon EMR son efímeros: los volúmenes se eliminan al terminar el clúster y las instancias (por ejemplo, al reducir grupos de instancias), por lo que no debe esperar que los datos persistan. Aunque los datos son efímeros, es posible que los datos en HDFS se puedan replicar en función del número y de la especialización de los nodos del clúster. Al agregar volúmenes de almacenamiento de Amazon EBS, estos se montan como volúmenes adicionales. No forman parte del volumen raíz. YARN está configurado para utilizar todos los volúmenes adicionales, pero usted es responsable de asignar los volúmenes adicionales como almacenamiento local (por ejemplo, para archivos de registro locales).

Consideraciones

Tenga en cuenta estas consideraciones adicionales cuando utilice Amazon EBS con clústeres de EMR:

- No puede realizar una instantánea de un volumen de Amazon EBS y posteriormente restaurarlo en Amazon EMR. Para crear configuraciones personalizadas reutilizables, utilice una AMI personalizada (disponible en Amazon EMR versión 5.7.0 y posteriores). Para obtener más información, consulte Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR.

- Solo se admite un volumen de dispositivo raíz cifrado de Amazon EBS cuando se utiliza una AMI personalizada. Para obtener más información, consulte Creación de una AMI personalizada con un volumen de dispositivo raíz de Amazon EBS cifrado.
- Si aplica etiquetas con la API de Amazon EMR, dichas operaciones se aplicarán a volúmenes de EBS.
- Existe un límite de 25 volúmenes por instancia.
- Los volúmenes de Amazon EBS en los nodos principales no pueden ser inferiores a 5 GB.
- Amazon EBS tiene un límite fijo de 2500 volúmenes de EBS por solicitud de lanzamiento de instancia. Este límite también se aplica a Amazon EMR en EC2 los clústeres. Le recomendamos que lance clústeres con el número total de volúmenes de EBS dentro de este límite y, a continuación, escale verticalmente el clúster de forma manual o con el escalado administrado por Amazon EMR, según sea necesario. Para obtener más información sobre el límite de volumen de EBS, consulte Cuotas de servicio.

Almacenamiento predeterminado de Amazon EBS para instancias

Para EC2 las instancias que tienen almacenamiento exclusivo para EBS, Amazon EMR asigna los volúmenes de almacenamiento gp2 o gp3 de Amazon EBS a las instancias. Al crear un clúster con la versión 5.22.0 y posteriores de Amazon EMR, la cantidad predeterminada de almacenamiento de Amazon EBS aumenta en función del tamaño de la instancia.

Dividimos el aumento del almacenamiento en varios volúmenes. Esto ofrece un mayor rendimiento de IOPS y, a su vez, un mejor rendimiento para algunas cargas de trabajo estandarizadas. Si desea utilizar una configuración diferente de almacenamiento para instancias de Amazon EBS, puede especificarla al crear un clúster de EMR o agregar nodos a un clúster existente. Puede utilizar los volúmenes gp2 o gp3 de Amazon EBS como volúmenes raíz y agregar volúmenes gp2 o gp3 como volúmenes adicionales. Para obtener más información, consulte Especificación de volúmenes de almacenamiento adicionales de EBS.

En la siguiente tabla se identifica el número predeterminado de volúmenes de almacenamiento gp2 de Amazon EBS, los tamaños y los tamaños totales por tipo de instancia. Para obtener información sobre los volúmenes gp2 en comparación con los gp3, consulte Comparación de los tipos de volúmenes gp2 y gp3 de Amazon EBS.

Volúmenes de almacenamiento gp2 de Amazon EBS predeterminados y tamaño por tipo de instancia para la versión 5.22.0 y posteriores de Amazon EMR

Tamaño de instancia	Número de volúmenes	Tamaño del volumen (GiB)	Tamaño total (GiB)
*.large	1	32	32
*.xlarge	2	32	64
*.2xlarge	4	32	128
*.4xlarge	4	64	256
*.8xlarge	4	128	512
9xlarge	4	144	576
*.10xlarge	4	160	640
*.12xlarge	4	192	768
*.16xlarge	4	256	1024
*.18xlarge	4	288	1152
*.24xlarge	4	384	1536

Volumen raíz de Amazon EBS predeterminado para las instancias

Con las versiones 6.15 y posteriores de Amazon EMR, Amazon EMR conecta automáticamente una SSD de uso general (gp3) de Amazon EBS como dispositivo raíz para mejorar el rendimiento. AMIs Con versiones anteriores, Amazon EMR adjunta el volumen SSD de uso general (gp2) de EBS como dispositivo raíz.

	6.15 y posteriores	6.14 y anteriores
Tipo de volumen raíz predeterminado		

	6.15 y posteriores	6.14 y anteriores
Tamaño predeterminado		
IOPS predeterminadas		
Rendimiento predeterminado		

Para obtener información sobre cómo personalizar el volumen del dispositivo raíz de Amazon EBS, consulte Especificación de volúmenes de almacenamiento adicionales de EBS.

Especificación de volúmenes de almacenamiento adicionales de EBS

Al configurar los tipos de instancia en Amazon EMR, puede especificar volúmenes de EBS adicionales, lo que agrega capacidad más allá del almacén de instancias (en caso de incluirse) y el volumen de EBS predeterminado. Amazon EBS ofrece los siguientes tipos de volúmenes: uso general (SSD), de IOPS aprovisionadas (SSD), de rendimiento optimizado (HDD), en frío (HDD) y magnéticos. Se diferencian en las características de rendimiento y en el precio, por lo que puede personalizar su almacenamiento en función de las necesidades empresariales y de análisis de sus aplicaciones. Por ejemplo, algunas aplicaciones pueden tener la necesidad de volcar contenido en el disco, mientras que otras pueden trabajar de forma segura en la memoria o con Amazon S3.

Solo puede asociar volúmenes de Amazon EBS a instancias durante el tiempo de inicio del clúster y cuando agrega un grupo de instancias de nodos de tarea adicional. Si una instancia en un clúster de Amazon EMR presenta errores, tanto la instancia como los volúmenes de Amazon EBS asociados se sustituirán con volúmenes nuevos. Por lo tanto, si separa manualmente un volumen de Amazon EBS, Amazon EMR lo trata como un error y sustituye tanto los almacenamientos de la instancia (si procede) como los almacenes de volumen.

Amazon EMR no le permite modificar el tipo de volumen de gp2 a gp3 para un clúster de EMR existente. Para utilizar gp3 en sus cargas de trabajo, lance un nuevo clúster de EMR. Además, no se recomienda actualizar el rendimiento y las IOPS de un clúster que esté en uso o que se esté aprovisionando, ya que Amazon EMR utiliza los valores de rendimiento y de IOPS que se especifican en el momento de lanzar el clúster para cualquier instancia nueva que agrega durante el escalado vertical del clúster. Para obtener más información, consulte Comparación de los tipos de volúmenes gp2 y gp3 de Amazon EBS y Selección de las IOPS y el rendimiento al migrar a tipos de volúmenes de Amazon EBS de gp3.

> **Important**
>
> Para utilizar un volumen gp3 con su clúster de EMR, lance un nuevo clúster.

Comparación de los tipos de volúmenes gp2 y gp3 de Amazon EBS

A continuación, se incluye una comparación del costo entre los volúmenes gp2 y gp3 en la región Este de EE UU. (Norte de Virginia). Para obtener la información más actualizada, consulte la página del producto Volúmenes de uso general de Amazon EBS y la Página de precios de Amazon EBS.

Tipo de volumen	gp3	gp2
Tamaño del volumen	1 GiB – 16 TiB	1 GiB – 16 TiB
IOPS predeterminadas o de referencia	3 000	De 3 IOPS/GiB (mínimo 100 IOPS) a un máximo de 16 000 IOPS. Los volúmenes de menos de 1 TiB también pueden alcanzar ráfagas de hasta 3000 IOPS.
Máximo de IOPS por volumen	16,000	16,000
Rendimiento predeterminado o de referencia	125 MiB/s	El límite de rendimiento está entre 128 MiB/s y 250 MiB/s, según el tamaño del volumen.
Rendimiento máximo por volumen	1000 MiB/s	250 MiB/s
Precio	0,08 USD al mes, 3000 IOPS gratuitas y 0,005 USD al mes por IOPS aprovisionadas, más de 3000 USD; 125 gratuitas y 0,04 USD aprovisionadas MiB/s MiB/s-month over 125MiB/s	0,10 USD/GiB al mes

Selección de las IOPS y el rendimiento al migrar a tipos de volúmenes de Amazon EBS de gp3

Al aprovisionar un volumen de gp2, debe calcular el tamaño del volumen para obtener las IOPS y el rendimiento proporcionales. Con gp3, no es necesario aprovisionar un volumen mayor para obtener un mayor rendimiento. Puede elegir el tamaño y el rendimiento que desee según las necesidades de la aplicación. Si selecciona el tamaño y los parámetros de rendimiento adecuados (IOPS, rendimiento), podrá reducir al máximo los costos sin que ello afecte al rendimiento.

A continuación se incluye una tabla que lo ayudará a seleccionar las opciones de configuración de gp3:

Tamaño del volumen	IOPS	Rendimiento
1–170 GiB	3 000	125 MiB/s
170–334 GiB	3 000	125 si MiB/s el tipo EC2 de instancia elegido admite MiB/s or less, use higher as per usage, Max 250 MiB/s 125 *.
334–1000 GiB	3 000	125 MiB/s si el tipo de EC2 instancia elegido admite 125 MiB/s or less, Use higher as per usage, Max 250 MiB/s *.
1000 GiB	Haga coincidir las IOPS de gp2 (tamaño en GiB x 3) o las IOPS máximas basadas en el volumen gp2 actual	125 MiB/s si el tipo de EC2 instancia elegido admite 125 MiB/s or less, Use higher as per usage, Max 250 MiB/s *.

*Gp3 tiene la capacidad de proporcionar un rendimiento de hasta 1000MiB/s. Since gp2 provides a maximum of 250MiB/s. Es posible que no necesite superar este límite cuando utilice gp3.

Configuración de redes en una VPC para Amazon EMR

La mayoría de los clústeres se lanzan en una red virtual mediante Amazon Virtual Private Cloud (Amazon VPC). Una VPC es una red virtual aislada AWS que está aislada de forma lógica dentro de su cuenta. AWS Puede configurar aspectos como los intervalos de direcciones IP privadas, las

subredes, las tablas de enrutamiento y las puertas de enlace de red. Para obtener más información, consulte la Guía del usuario de Amazon VPC.

VPC ofrece las siguientes capacidades:

- Procesamiento de información confidencial

  Lanzar un clúster en una VPC es similar a lanzar el clúster en una red privada con herramientas adicionales, como tablas de enrutamiento y redesACLs, para definir quién tiene acceso a la red. Si va a procesar información confidencial en su clúster, es posible que desee el control de acceso adicional que ofrece el lanzamiento del clúster en una VPC. Además, puede elegir lanzar sus recursos en una subred privada en la que ninguno de ellos tenga conectividad directa a Internet.
- Acceso a los recursos de una red interna

  Si la fuente de datos se encuentra en una red privada, puede resultar poco práctico o indeseable cargar esos datos AWS para importarlos a Amazon EMR, ya sea por la cantidad de datos que se van a transferir o por la naturaleza confidencial de los datos. En su lugar, puede lanzar el clúster en una VPC y conectar su centro de datos a su VPC a través de una conexión de VPN, habilitando el clúster para que acceda a recursos de su red interna. Por ejemplo, si tiene una base de datos Oracle en su centro de datos, el lanzamiento del clúster en una VPC conectada a dicha red por VPN permite que el clúster acceda a la base de datos Oracle.

Subredes públicas y privadas

Puede lanzar clústeres de Amazon EMR tanto en subredes de VPC públicas y privadas. Esto significa que no necesita conectividad a Internet para ejecutar un clúster de Amazon EMR; sin embargo, es posible que deba configurar la traducción de direcciones de red (NAT) y las puertas de enlace VPN para acceder a los servicios o recursos ubicados fuera de la VPC, por ejemplo, en una intranet corporativa o en puntos finales de servicio público como. AWS AWS Key Management Service

> **Important**
>
> Amazon EMR solo es compatible con el lanzamiento de clústeres en subredes privadas en su versión 4.2 y posteriores.

Para obtener más información sobre Amazon VPC, consulte la Guía del usuario de Amazon VPC.

Temas

- Opciones de Amazon VPC al lanzar un clúster
- Configuración de una VPC para alojar clústeres de Amazon EMR
- Lanzamiento de clústeres en una VPC con Amazon EMR
- Ejemplos de políticas para subredes privadas que acceden a Amazon S3
- Más recursos para obtener más información sobre VPCs

Opciones de Amazon VPC al lanzar un clúster

Cuando se lanza un clúster de Amazon EMR en una VPC, puede lanzarse en una subred compartida, privada o pública. Existen ligeras pero importantes diferencias en la configuración, en función del tipo de subred que elija para un clúster.

Subredes públicas

Los clústeres de EMR de una subred pública requieren una gateway de Internet conectada. Esto se debe a que los clústeres de Amazon EMR deben acceder a los AWS servicios y a Amazon EMR. Si un servicio, como Amazon S3, ofrece la posibilidad de crear un punto de conexión de VPC, puede acceder a dichos servicios mediante el punto de conexión, en lugar de acceder a un punto de conexión público a través de una puerta de enlace de Internet. Además, Amazon EMR no puede comunicarse con clústeres en subredes públicas a través de un dispositivo de traducción de direcciones de red (NAT). Para este fin se necesita una gateway de Internet, pero puede utilizar una instancia NAT o una gateway para otro tráfico en situaciones más complejas.

Todas las instancias de un clúster se conectan a Amazon S3 a través de un punto de conexión de VPC o una puerta de enlace de Internet. Otros AWS servicios que actualmente no admiten puntos finales de VPC utilizan solo una puerta de enlace a Internet.

Si tiene AWS recursos adicionales que no desea conectar a la puerta de enlace de Internet, puede lanzar esos componentes en una subred privada que cree dentro de su VPC.

Los clústeres que se ejecutan en una subred pública utilizan dos grupos de seguridad: uno para el nodo principal y otro para los nodos básicos y de tarea. Para obtener más información, consulte Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

El siguiente diagrama muestra cómo se ejecuta un clúster de Amazon EMR en una VPC mediante una subred pública. El clúster puede conectarse a otros AWS recursos, como los buckets de Amazon S3, a través de la puerta de enlace de Internet.

En el siguiente diagrama muestra cómo configurar una VPC para que un clúster en la VPC pueda acceder a los recursos en su propia red, como, por ejemplo, una base de datos Oracle.

Subredes privadas

Una subred privada le permite lanzar AWS recursos sin necesidad de que la subred tenga una puerta de enlace de Internet conectada. Amazon EMR es compatible con el lanzamiento de clústeres en subredes privadas en su versión 4.2.0 y posteriores.

> **Note**
>
> Al configurar un clúster de Amazon EMR en una subred privada, se recomienda que también configure puntos de conexión de VPC para Amazon S3. Si su clúster de EMR se encuentra en una subred privada sin puntos de conexión de VPC para Amazon S3, incurrirá en cargos adicionales de puerta de enlace de NAT asociados al tráfico de S3, ya que el tráfico entre su clúster de EMR y S3 no permanecerá dentro de su VPC.

Las subredes privadas se diferencian de las públicas en los siguientes aspectos:

- Para acceder a AWS los servicios que no proporcionan un punto de enlace de VPC, debe seguir utilizando una instancia de NAT o una puerta de enlace de Internet.
- Como mínimo, deberá proporcionar una ruta al bucket de registros de servicio de Amazon EMR y al repositorio de Amazon Linux en Amazon S3. Para obtener más información, consulte Ejemplos de políticas para subredes privadas que acceden a Amazon S3
- Si utiliza las características de EMRFS, debe disponer de un punto de conexión de VPC de Amazon S3 y de una ruta desde la subred privada a DynamoDB.
- La depuración solo funciona si proporciona una ruta desde la subred privada a un punto de conexión de Amazon SQS público.
- La creación de una configuración de subred privada con una instancia de NAT o puerto de enlace en una subred pública solo es compatible con la AWS Management Console. La forma más sencilla de agregar y configurar instancias NAT y puntos de conexión de VPC de Amazon S3 para los clústeres de Amazon EMR consiste en utilizar la página Lista de subredes de la VPC de la consola de Amazon EMR. Para configurar puertas de enlace de NAT, consulte Puertas de enlace de NAT en la Guía del usuario de Amazon VPC.

- No puede cambiar una subred con un clúster de Amazon EMR existente de pública a privada o viceversa. Para localizar un clúster de Amazon EMR dentro de una subred privada, el clúster debe iniciarse en dicha subred privada.

Amazon EMR crea y utiliza diferentes grupos de seguridad predeterminados para los clústeres de una subred privada: ElasticMapReduce-Master-Private, ElasticMapReduce -Slave-Private y -. ElasticMapReduce ServiceAccess Para obtener más información, consulte Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

Para obtener una lista completa NACLs de su clúster, elija Grupos de seguridad para Primary y Security Groups for Core & Task en la página de detalles del clúster de la consola Amazon EMR.

La imagen siguiente muestra cómo se configura un clúster de Amazon EMR dentro de una subred privada. La única comunicación fuera de la subred es a Amazon EMR.

La imagen siguiente muestra una configuración de ejemplo para un clúster de Amazon EMR dentro de una subred privada conectada a una instancia de NAT que reside en una subred pública.

Subredes compartidas

El uso compartido de VPC permite a los clientes compartir subredes con otras AWS cuentas de la misma organización. AWS Puede lanzar clústeres de Amazon EMR en subredes compartidas públicas o privadas, con las siguientes advertencias.

El propietario de la subred debe compartir una subred con usted antes de lanzar un clúster de Amazon EMR en ella. Sin embargo, las subredes compartidas se pueden dejar de compartir más adelante. Para obtener más información, consulte Trabajar con Shared. VPCs Cuando un clúster se lanza en una subred compartida y esa subred se deja de compartir, es posible que observe comportamientos específicos basados en el estado del clúster de Amazon EMR cuando la subred se deja de compartir.

- La subred se deja de compartir antes de que el clúster se haya lanzado correctamente: si el propietario deja de compartir la VPC o la subred de Amazon mientras que el participante se está lanzando un clúster, es posible que el clúster no se inicie o que se inicialice parcialmente sin aprovisionar todas las instancias solicitadas.
- La subred se deja de compartir después de que el clúster se haya lanzado correctamente: cuando el propietario deja de compartir una subred o VPC de Amazon con el participante, los clústeres del

participante no podrán cambiar de tamaño para añadir nuevas instancias o para sustituir instancias en mal estado.

Cuando se lanza un clúster de Amazon EMR, se crean varios grupos de seguridad. En una subred compartida, el participante de la subred controla estos grupos de seguridad. El propietario de la subred pueden ver estos grupos de seguridad, pero no puede realizar ninguna acción en ellos. Si el propietario de la subred quiere eliminar o modificar el grupo de seguridad, el participante que ha creado el grupo de seguridad debe realizar la acción.

Control de permisos de VPC con IAM

De forma predeterminada, todos los usuarios de pueden ver todas las subredes de la cuenta y cualquier usuario puede lanzar un clúster en cualquier subred.

Cuando lanza un clúster en una VPC, puede usar AWS Identity and Access Management (IAM) para controlar el acceso a los clústeres y restringir las acciones mediante políticas, tal como lo haría con los clústeres lanzados en Amazon Classic. EC2 Para más información acerca de IAM, consulte la Guía del usuario de IAM.

También puede utilizar IAM para controlar quién puede crear y administrar subredes. Por ejemplo, puede crear un rol de IAM para administrar subredes y un segundo rol que pueda lanzar clústeres, pero no pueda modificar la configuración de Amazon VPC. Para obtener más información sobre la administración de políticas y acciones en Amazon EC2 y Amazon VPC, consulte Políticas de IAM para Amazon en la Guía EC2 del usuario de Amazon EC2 .

Configuración de una VPC para alojar clústeres de Amazon EMR

Antes de poder lanzar clústeres en una VPC, debe crear una VPC y una subred. Para subredes públicas, debe crear una gateway de Internet y asociarla a la subred. Las siguientes instrucciones describen cómo crear una VPC capaz de alojar clústeres de Amazon EMR.

Para crear una VPC con subredes para un clúster de Amazon EMR

1. Abra la consola de Amazon VPC en https://console.aws.amazon.com/vpc/.
2. En la parte superior derecha de la página, elija la Región de AWS para su VPC.
3. Seleccione Creación de VPC.
4. En la página Crear VPC, seleccione VPC y más.

5. En Generación automática de etiquetas de nombre, habilite la generación automática e ingrese un nombre para la VPC. Esto le ayudará a identificar la VPC y la subred en la consola de Amazon VPC después de crearlas.
6. En el campo de bloqueo IPv4 CIDR, introduzca un espacio de direcciones IP privadas para su VPC a fin de garantizar una resolución adecuada del nombre de host DNS; de lo contrario, puede que se produzcan errores en el clúster de Amazon EMR. Esto incluye los siguientes rangos de direcciones IP:
   - 10.0.0.0 - 10.255.255.255
   - 172.16.0.0 - 172.31.255.255
   - 192.168.0.0 - 192.168.255.255
7. En Número de zonas de disponibilidad (AZs), elija el número de zonas de disponibilidad en las que desea lanzar las subredes.
8. En Número de subredes públicas, elija una sola subred pública para agregar a la VPC. Si los datos que usa el clúster están disponibles en Internet (por ejemplo, en Amazon S3 o Amazon RDS), solo necesita utilizar una subred pública y no necesita agregar una subred privada.
9. En Number of private subnets (Número de subredes privadas), elija el número de subredes públicas que desea agregar a la VPC. Seleccione una o más si los datos de la aplicación se almacenan en su propia red (por ejemplo, en una base de datos de Oracle). Para una VPC en una subred privada, todas las EC2 instancias de Amazon deben tener como mínimo una ruta a Amazon EMR a través de la elastic network interface. En la consola, esto se configura automáticamente.
10. De manera opcional, en Puertas de enlace de NAT, puede optar por agregar puertas de enlace de NAT. Solo son necesarias si tiene subredes privadas que necesitan comunicarse con Internet.
11. En Puntos de conexión de la VPC, elija de manera opcional agregar puntos de conexión para Amazon S3 en sus subredes.
12. Compruebe que las opciones Habilitar nombres de host DNS y Habilitar la resolución de DNS estén marcadas. Para obtener más información, consulte Utilización de DNS con su VPC.
13. Seleccione Creación de VPC.
14. En una ventana de estado se muestra el trabajo en curso. Cuando finalice el trabajo, elija Ver VPC para ir a la VPCs página Tu, que muestra tu VPC predeterminada y la VPC que acabas de crear. La VPC que acaba de crear es una VPC no predeterminada. Por lo tanto, la columna Default VPC muestra No.

15. Si desea asociar su VPC a una entrada de DNS que no incluya un nombre de dominio, vaya a Conjuntos de opciones de DHCP, elija Crear conjunto de opciones de DHCP y omita un nombre de dominio. Después de crear el conjunto de opciones, vaya a la nueva VPC, elija Editar conjunto de opciones de DHCP en el menú Acciones y seleccione el nuevo conjunto de opciones. No puede editar el nombre de dominio utilizando la consola después de haber creado el conjunto de opciones de DNS.

    Es una práctica recomendada con Hadoop y aplicaciones relacionadas garantizar la resolución del nombre de dominio completo (FQDN) para los nodos. Para garantizar una resolución de DNS adecuada, configure una VPC que incluya un conjunto de opciones de DHCP cuyos parámetros se definan con los valores siguientes:

    - domain-name = **`ec2.internal`**

      Use **`ec2.internal`** si su región es Este de EE. UU. (Norte de Virginia). Para otras regiones, utilice. *`region-name`* **`.compute.internal`** Para ver ejemplos de `us-west-2`, utilice **`us-west-2.compute.internal`**. Para la región AWS GovCloud (EE. UU.-Oeste), utilice**`us-gov-west-1.compute.internal`**.
    - domain-name-servers = **`AmazonProvidedDNS`**

    Para más información, consulte Conjuntos de opciones de DHCP en la Guía del usuario de Amazon VPC.

16. Después de haber creado la VPC, vaya a la página Subredes y anote el ID de subred de una de las subredes de su nueva VPC. Utilice esta información al lanzar el clúster de Amazon EMR en la VPC.

Lanzamiento de clústeres en una VPC con Amazon EMR

Después de tener una subred configurada para alojar clústeres de Amazon EMR, lance el clúster en dicha subred especificando el identificador de subred asociado al crear el clúster.

> **Note**
>
> Amazon EMR admite subredes privadas en su versión 4.2 y superiores.

Cuando se lanza el clúster, Amazon EMR agrega grupos de seguridad en función de si el clúster se lanza en subredes privadas o públicas de VPC. Todos los grupos de seguridad permiten la entrada en el puerto 8443 para comunicarse con el servicio de Amazon EMR, pero los intervalos de direcciones IP varían en subredes públicas y privadas. Amazon EMR administra todos estos grupos de seguridad y, con el tiempo, es posible que necesite añadir direcciones IP adicionales al AWS rango. Para obtener más información, consulte Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

Para administrar el clúster en una VPC, Amazon EMR asocia un dispositivo de red al nodo principal y lo administra a través de este dispositivo. Puedes ver este dispositivo mediante la acción de la EC2 API de Amazon `DescribeInstances`. Si modifica este dispositivo de algún modo, el clúster podría fallar.

Console

Para lanzar un clúster en una VPC utilizando la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Redes, vaya al campo Nube privada virtual (VPC). Ingrese el nombre de la VPC o elija Buscar para seleccionar la VPC. También puede elegir Crear VPC para crear una VPC que pueda utilizar para su clúster.
4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para lanzar un clúster en una VPC con AWS CLI

> **Note**
>
> AWS CLI No proporciona una forma de crear una instancia de NAT automáticamente y conectarla a su subred privada. Sin embargo, para crear un punto de conexión de S3 en su subred, puede utilizar los comandos de la CLI de Amazon VPC. Utilice la consola para crear instancias NAT y lanzar clústeres en una subred privada.

Una vez que la VPC se ha configurado, puede lanzar clústeres de Amazon EMR en ella utilizando el subcomando `create-cluster` con el parámetro `--ec2-attributes`. Utilice el parámetro `--ec2-attributes` para especificar la subred de VPC para el clúster.

- Para crear un clúster en una subred específica, escribe el siguiente comando, sustitúyelo por *myKey* el nombre de tu EC2 key pair de Amazon y *77XXXX03* sustitúyelo por tu ID de subred.

```
aws emr create-cluster --name "Test cluster" --release-label emr-4.2.0 --
applications Name=Hadoop Name=Hive Name=Pig --use-default-roles --ec2-attributes
 KeyName=myKey,SubnetId=subnet-77XXXX03 --instance-type m5.xlarge --instance-
count 3
```

  Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos básicos. Todos los nodos utilizan el tipo de instancia especificado en el comando.

> **Note**
>
> Si no ha creado anteriormente el rol de servicio Amazon EMR y el perfil de EC2 instancia predeterminados, escriba `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Garantizar la disponibilidad de las direcciones IP para un clúster de EMR en EC2

Para garantizar que haya disponible una subred con suficientes direcciones IP libres en el momento del lanzamiento, la selección de EC2 subredes comprueba la disponibilidad de IP. Si el proceso de creación utiliza una subred con el número de direcciones IP necesario para lanzar los nodos básicos, principales y de tareas según sea necesario, incluso si, tras la creación inicial, solo se crean los nodos básicos del clúster. El EMR comprueba la cantidad de direcciones IP necesarias para lanzar los nodos principales y de tareas durante la creación, y calcula por separado la cantidad de direcciones IP necesarias para lanzar los nodos básicos. Amazon EMR determina automáticamente el número mínimo de instancias o nodos principales y de tareas necesarios.

> **Important**
>
> Si ninguna subred de la VPC tiene suficientes subredes IPs disponibles para alojar los nodos esenciales, se devuelve un error y no se crea el clúster.

En la mayoría de los casos de implementación, hay una diferencia de tiempo entre cada lanzamiento de los nodos principales, básicos y de tareas. Además, es posible que varios clústeres compartan una subred. En estos casos, la disponibilidad de las direcciones IP puede fluctuar y, por ejemplo, los lanzamientos posteriores de nodos de tareas pueden estar limitados por las direcciones IP disponibles.

Ejemplos de políticas para subredes privadas que acceden a Amazon S3

En el caso de las subredes privadas, como mínimo debe proporcionar a Amazon EMR la capacidad de acceder a los repositorios de Amazon Linux. Esta política de subred privada forma parte de las políticas de puntos de conexión de VPC para el acceso a Amazon S3.

Con Amazon EMR 5.25.0 o posterior, para habilitar el acceso de un clic al servidor del historial de Spark persistente, debe permitir a Amazon EMR el acceso al bucket del sistema que recopila los registros de eventos de Spark. Si habilitas el registro, proporciona permisos PUT al siguiente depósito:

```
aws157-logs-${AWS::Region}/*
```

Para más información, consulte Acceso de un clic al servidor del historial de Spark persistente.

Usted debe determinar las restricciones de política que satisfacen sus necesidades empresariales. En el siguiente ejemplo de política se proporcionan permisos para acceder a los repositorios de Amazon Linux y al bucket del sistema de Amazon EMR para recopilar registros de eventos de Spark. Muestra algunos ejemplos de nombres de recursos para los buckets.

Para más información acerca del uso de políticas de IAM con puntos de conexión de Amazon VPC, consulte Políticas de punto de conexión para Amazon S3.

En el siguiente ejemplo de política se describen ejemplos de recursos en la región us-east-1.

JSON

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AmazonLinuxAMIRepositoryAccess",
      "Effect": "Allow",
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::packages.us-east-1.amazonaws.com/*",
        "arn:aws:s3:::repo.us-east-1.amazonaws.com/*"
      ]
    },
    {
      "Sid": "EnableApplicationHistory",
      "Effect": "Allow",
      "Action": [
        "s3:Put*",
        "s3:Get*",
        "s3:Create*",
        "s3:Abort*",
        "s3:List*"
      ],
      "Resource": [
        "arn:aws:s3:::prod.us-east-1.appinfo.src/*"
      ]
    }
  ]
}
```

El siguiente ejemplo de política proporciona los permisos necesarios para acceder a los repositorios de Amazon Linux 2 en la región us-east-1.

```
{
   "Statement": [
       {
           "Sid": "AmazonLinux2AMIRepositoryAccess",
           "Effect": "Allow",
           "Principal": "*",
           "Action": "s3:GetObject",
           "Resource": [
            "arn:aws:s3:::amazonlinux.us-east-1.amazonaws.com/*",
```

```
            "arn:aws:s3:::amazonlinux-2-repos-us-east-1/*"
           ]
       }
   ]
}
```

El siguiente ejemplo de política proporciona los permisos necesarios para acceder a los repositorios de Amazon Linux 2023 en la región us-east-1.

```
{
    "Statement": [
        {
            "Sid": "AmazonLinux2023AMIRepositoryAccess",
            "Effect": "Allow",
            "Principal": "*",
            "Action": "s3:GetObject",
            "Resource": [
                 "arn:aws:s3:::al2023-repos-us-east-1-de612dc2/*"
            ]
        }
    ]
 }
```

Regiones disponibles

La siguiente tabla contiene una lista de buckets por región e incluye un nombre de recurso de Amazon (ARN) para el repositorio, así como una cadena que representa el ARN para el `appinfo.src`. El ARN, o nombre de recurso de Amazon, es una cadena que identifica un recurso de forma exclusiva. AWS

Región	Buckets de repositorio	AppInfo bucket
EE.UU. Este (Ohio)	"arn:aws:s3:::packages.us-east-2.amazonaws.com/","arn:aws:s3:::repo.us-east-2.amazonaws.com/","arn:aws:s3:::repo.us-east-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.us-east-2.appinfo.src/*"
EE.UU. Este (Norte de Virginia)	"arn:aws:s3:::packages.us-east-1.amazonaws.com/","arn:aws:s3:::repo.us-east-1.amazonaws.co	"arn:aws:s3:::prod.us-east-1.appinfo.src/*"

Región	Buckets de repositorio	AppInfo bucket
	m/","arn:aws:s3:::repo.us-east-1.emr.amazonaws.com/*"	
EE.UU. Oeste (Norte de California)	"arn:aws:s3:::packages.us-west-1.amazonaws.com/","arn:aws:s3:::repo.us-west-1.amazonaws.com/","arn:aws:s3:::repo.us-west-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.us-west-1.appinfo.src/*"
EE.UU. Oeste (Oregón)	"arn:aws:s3:::packages.us-west-2.amazonaws.com/","arn:aws:s3:::repo.us-west-2.amazonaws.com/","arn:aws:s3:::repo.us-west-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.us-west-2.appinfo.src/*"
África (Ciudad del Cabo)	"arn:aws:s3:::packages.af-south-1.amazonaws.com/","arn:aws:s3:::repo.af-south-1.amazonaws.com/","arn:aws:s3:::repo.af-south-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.af-south-1.appinfo.src/*"
África (Ciudad del Cabo)	"arn:aws:s3:::packages.ap-east-1.amazonaws.com/","arn:aws:s3:::repo.ap-east-1.amazonaws.com/","arn:aws:s3:::repo.ap-east-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-east-1.appinfo.src/*"
Asia-Pacífico (Hyderabad)	"arn:aws:s3:::packages.ap-south-2.amazonaws.com/","arn:aws:s3:::repo.ap-south-2.amazonaws.com/","arn:aws:s3:::repo.ap-south-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-south-2.appinfo.src/*"

Región	Buckets de repositorio	AppInfo bucket
Asia-Pacífico (Yakarta)	"arn:aws:s3:::packages.ap-southeast-3.amazonaws.com/", "arn:aws:s3:::repo.ap-southeast-3.amazonaws.com/","arn:aws:s3:::repo.ap-southeast-3.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-southeast-3.appinfo.src/*"
Asia-Pacífico (Malasia)	"arn:aws:s3:::packages.ap-southeast-5.amazonaws.com/", "arn:aws:s3:::repo.ap-southeast-5.amazonaws.com/","arn:aws:s3:::repo.ap-southeast-5.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-southeast-5.appinfo.src/*"
Asia-Pacífico (Melbourne)	"arn:aws:s3:::packages.ap-southeast-4.amazonaws.com/", "arn:aws:s3:::repo.ap-southeast-4.amazonaws.com/","arn:aws:s3:::repo.ap-southeast-4.emr.amazonaws.com/*"	«arn:aws:s3: ::prod.ap-south-4. appinfo.src/*»
Asia-Pacífico (Mumbai)	"arn:aws:s3:::packages.ap-south-1.amazonaws.com/","arn:aws:s3:::repo.ap-south-1.amazonaws.com/","arn:aws:s3: ::repo.ap-south-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-south-1.appinfo.src/*"
Asia-Pacífico (Osaka)	«arn:aws:s3::: packages.ap-northeast-3.amazonaws.com/», «arn:aws: s3::: repo.ap-northeast-3.amazonaws.com/», «arn:aws:s3: ::repo.ap-northeast-3.emr.amazonaws.com/*»	«arn:aws:s3: :prod.ap-northeast -3.appinfo.src/*»
Asia Pacífico (Seúl)	"arn:aws:s3:::packages.ap-northeast-2.amazonaws.com/", "arn:aws:s3:::repo.ap-northeast-2.amazonaws.com/","arn:aws:s3:::repo.ap-northeast-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-northeast-2.appinfo.src/*"

Región	Buckets de repositorio	AppInfo bucket
Asia-Pacífico (Singapur)	"arn:aws:s3:::packages.ap-southeast-1.amazonaws.com/", "arn:aws:s3:::repo.ap-southeast-1.amazonaws.com/","arn:aws:s3:::repo.ap-southeast-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-southeast-1.appinfo.src/*"
Asia Pacífico (Sídney)	"arn:aws:s3:::packages.ap-southeast-2.amazonaws.com/", "arn:aws:s3:::repo.ap-southeast-2.amazonaws.com/","arn:aws:s3:::repo.ap-southeast-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-southeast-2.appinfo.src/*"
Asia Pacífico (Tokio)	"arn:aws:s3:::packages.ap-northeast-1.amazonaws.com/", "arn:aws:s3:::repo.ap-northeast-1.amazonaws.com/","arn:aws:s3:::repo.ap-northeast-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ap-northeast-1.appinfo.src/*"
Canadá (Central)	"arn:aws:s3:::packages.ca-central-1.amazonaws.com/","arn:aws:s3:::repo.ca-central-1.amazonaws.com/","arn:aws:s3:::repo.ca-central-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.ca-central-1.appinfo.src/*"
Oeste de Canadá (Calgary	«arn:aws:s3::: packages.ca-west-1.amazonaws.com/», «arn:aws:s3:::repo.ca-west-1.amazonaws.com/», «arn:aws:s3: ::repo.ca-west-1.emr.amazonaws.com/*»	«arn:aws:s3: ::prod.ca-west-1.appinfo.src/*»
Europa (Fráncfort)	"arn:aws:s3:::packages.eu-central-1.amazonaws.com/","arn:aws:s3:::repo.eu-central-1.amazonaws.com/","arn:aws:s3:::repo.eu-central-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-central-1.appinfo.src/*"

Región	Buckets de repositorio	AppInfo bucket
Europa (Irlanda)	"arn:aws:s3:::packages.eu-west-1.amazonaws.com/","arn:aws:s3:::repo.eu-west-1.amazonaws.com/","arn:aws:s3:::repo.eu-west-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-west-1.appinfo.src/*"
Europa (Londres)	"arn:aws:s3:::packages.eu-west-2.amazonaws.com/","arn:aws:s3:::repo.eu-west-2.amazonaws.com/","arn:aws:s3:::repo.eu-west-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-west-2.appinfo.src/*"
Europa (Milán)	"arn:aws:s3:::packages.eu-south-1.amazonaws.com/","arn:aws:s3:::repo.eu-south-1.amazonaws.com/","arn:aws:s3:::repo.eu-south-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-south-1.appinfo.src/*"
Europa (París)	"arn:aws:s3:::packages.eu-west-3.amazonaws.com/","arn:aws:s3:::repo.eu-west-3.amazonaws.com/","arn:aws:s3:::repo.eu-west-3.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-west-3.appinfo.src/*"
Europa (España)	"arn:aws:s3:::packages.eu-south-2.amazonaws.com/","arn:aws:s3:::repo.eu-south-2.amazonaws.com/","arn:aws:s3:::repo.eu-south-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-south-2.appinfo.src/*"

Región	Buckets de repositorio	AppInfo bucket
Europa (Estocolmo)	"arn:aws:s3:::packages.eu-north-1.amazonaws.com/","arn:aws:s3::repo.eu-north-1.amazonaws.com/","arn:aws:s3:::repo.eu-north-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-north-1.appinfo.src/*"
Europa (Zúrich)	"arn:aws:s3:::packages.eu-central-2.amazonaws.com/","arn:aws:s3:::repo.eu-central-2.amazonaws.com/","arn:aws:s3:::repo.eu-central-2.emr.amazonaws.com/*"	"arn:aws:s3:::prod.eu-central-2.appinfo.src/*"
Israel (Tel Aviv)	"arn:aws:s3:::packages.il-central-1.amazonaws.com/","arn:aws:s3:::repo.il-central-1.amazonaws.com/","arn:aws:s3:::repo.il-central-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.il-central-1.appinfo.src/*"
Medio Oriente (Baréin)	"arn:aws:s3:::packages.me-south-1.amazonaws.com/","arn:aws:s3:::repo.me-south-1.amazonaws.com/","arn:aws:s3::repo.me-south-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.me-south-1.appinfo.src/*"
Medio Oriente (EAU)	"arn:aws:s3:::packages.me-central-1.amazonaws.com/","arn:aws:s3:::repo.me-central-1.amazonaws.com/","arn:aws:s3:::repo.me-central-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.me-central-1.appinfo.src/*"
América del Sur (São Paulo)	"arn:aws:s3:::packages.sa-east-1.amazonaws.com/","arn:aws:s3:::repo.sa-east-1.amazonaws.com/","arn:aws:s3:::repo.sa-east-1.emr.amazonaws.com/*"	"arn:aws:s3:::prod.sa-east-1.appinfo.src/*"

Región	Buckets de repositorio	AppInfo bucket
AWS GovCloud (EEUU-Este)	«arn:aws:s3: ::paquetes. us-gov-east-1.amazonaws.com/», «arn:aws: s3: ::repo. us-gov-east-1.amazonaws.com/», «arn:aws:s3: ::repo. us-gov-east-1.emr.amazonaws.com/*»	«arn:aws:s3: ::prod. us-gov-east-1.appinfo.src/*»
AWS GovCloud (EE. UU.-Oeste)	«arn:aws:s3: ::paquetes. us-gov-west-1.amazonaws.com/», «arn:aws: s3: ::repo. us-gov-west-1.amazonaws.com/», «arn:aws:s3: ::repo. us-gov-west-1.emr.amazonaws.com/*»	"arn:aws:s3:::prod.me-south-1.appinfo.src/*"

Más recursos para obtener más información sobre VPCs

Utilice los siguientes temas para obtener más información sobre las subredes VPCs y las subredes.

- Subredes privadas en una VPC
  - Escenario 2: VPC con subredes públicas y privadas (NAT)
  - Instancias NAT
  - Alta disponibilidad para instancias NAT de Amazon VPC: un ejemplo
- Subredes públicas en una VPC
  - Escenario 1: VPC con una única subred pública
- Información de VPC general
  - Guía del usuario de Amazon VPC
  - Interconexión de VPC
  - Uso de interfaces de red elásticas con una VPC
  - Securely connect to Linux instances running in a private VPC

Creación de un clúster de Amazon EMR con flotas de instancias o grupos de instancias uniformes

Al crear un clúster y especificar la configuración del nodo principal y de los nodos básicos y de tareas, tiene dos opciones de configuración. Puede utilizar flotas de instancias o grupos de instancias

uniformes. La opción de configuración que elija se aplica a todos los nodos, se aplica durante toda la vida útil del clúster y las flotas de instancias y los grupos de instancias no pueden coexistir en un clúster. La configuración de las flotas de instancias está disponible en la versión 4.8.0 y posteriores de Amazon EMR, excluyendo las versiones 5.0.x.

Puede utilizar la consola de Amazon EMR AWS CLI, la o la API de Amazon EMR para crear clústeres con cualquier configuración. Al utilizar el comando `create-cluster` desde la AWS CLI, puede utilizar los parámetros `--instance-fleets` para crear el clúster utilizando las flotas de instancias o, como opción, utilizar los parámetros `--instance-groups` para crearlo usando grupos de instancias uniformes.

Lo mismo sucede utilizando la API de Amazon EMR. Puede utilizar la configuración `InstanceGroups` para especificar una matriz de objetos `InstanceGroupConfig` o bien utilizar la configuración `InstanceFleets` para especificar una matriz de objetos `InstanceFleetConfig`.

En la nueva consola de Amazon EMR, puede elegir utilizar grupos de instancias o flotas de instancias al crear un clúster y tiene la opción de utilizar instancias de spot con cada uno de ellos. En la consola anterior de Amazon EMR, si utiliza la configuración predeterminada de las Opciones rápidas al crear un clúster, Amazon EMR aplica la configuración de los grupos de instancias uniformes al clúster y utiliza instancias bajo demanda. Para utilizar las instancias de spot con grupos de instancias uniformes o para configurar las flotas de instancias y otras personalizaciones, elija Advanced Options (Opciones avanzadas).

Flotas de instancias

La configuración de las flotas de instancias ofrece la más amplia variedad de opciones de aprovisionamiento para las instancias de Amazon EC2 . Cada tipo de nodo tiene una sola flota de instancia y la flota de instancia de tarea es opcional. Puede especificar hasta cinco tipos de EC2 instancias por flota o 30 tipos de EC2 instancias por flota al crear un clúster mediante la AWS CLI API Amazon EMR y una estrategia de asignación para instancias puntuales y bajo demanda. En el caso de las flotas de instancia básica y de tarea, se asigna una capacidad de destino para instancias bajo demanda y otra para instancias de spot. Amazon EMR elige cualquier combinación de los tipos de instancia especificados para satisfacer las capacidades de destino, aprovisionando tanto las instancias bajo demanda como las instancias de spot.

En el caso del tipo de nodo principal, Amazon EMR elige un solo tipo de instancia de la lista de instancias y especifica si está aprovisionado como instancia bajo demanda o instancia de spot. Las flotas de instancias también ofrecen opciones adicionales para las compras de instancias de spot y bajo demanda. Las opciones de instancias de spot incluyen un tiempo de espera que especifica

la acción que se debe realizar en caso de que no se pueda aprovisionar la capacidad de spot y una estrategia de asignación preferida (optimizada en función de la capacidad) para lanzar flotas de instancias de spot. Las flotas de instancias bajo demanda también se pueden lanzar mediante la opción de estrategia de asignación (precio más bajo). Si utiliza un rol de servicio que no es el predeterminado de EMR o utiliza una política administrada de EMR en su rol de servicio, debe agregar permisos adicionales al rol de servicio de clúster personalizado para habilitar la opción de estrategia de asignación. Para obtener más información, consulte Rol de servicio para Amazon EMR (rol de EMR).

Para más información acerca de la configuración de las flotas de instancias, consulte Planificación y configuración de flotas de instancias para su clúster de Amazon EMR.

Grupos de instancias uniformes

Los grupos de instancias uniformes ofrecen una configuración más sencilla que las flotas de instancias. Cada clúster de Amazon EMR puede incluir hasta 50 grupos de instancias: un grupo de instancias principal que contiene una EC2 instancia de Amazon, un grupo de instancias principal que contiene una o más EC2 instancias y hasta 48 grupos de instancias de tareas opcionales. Cada grupo de instancias principales y de tareas puede contener cualquier número de EC2 instancias de Amazon. Puedes escalar cada grupo de instancias añadiendo y eliminando EC2 instancias de Amazon manualmente, o puedes configurar el escalado automático. Para obtener información sobre la adición y eliminación de instancias, consulte Utilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes.

Para obtener más información acerca de la configuración de grupos de instancias uniformes, consulte Configuración de grupos de instancias uniformes para su clúster de Amazon EMR.

Uso de flotas de instancias y grupos de instancias

Temas

- Planificación y configuración de flotas de instancias para su clúster de Amazon EMR
- Reconfiguración de las flotas de instancias para su clúster de Amazon EMR
- Uso de las reservas de capacidad con flotas de instancias en Amazon EMR
- Configuración de grupos de instancias uniformes para su clúster de Amazon EMR
- Flexibilidad de zona de disponibilidad para un clúster de Amazon EMR
- Configuración de tipos de instancias de clúster de Amazon EMR y prácticas recomendadas para instancias de Spot

Planificación y configuración de flotas de instancias para su clúster de Amazon EMR

> **Note**
>
> La configuración de las flotas de instancias está disponible solo en las versiones 4.8.0 y posteriores de Amazon EMR, excluidas las versiones 5.0.0 y 5.0.3.

La configuración de la flota de instancias para los clústeres de Amazon EMR le permite seleccionar una amplia variedad de opciones de aprovisionamiento para EC2 las instancias de Amazon y le ayuda a desarrollar una estrategia de recursos flexible y elástica para cada tipo de nodo del clúster.

En una configuración de flota de instancias, debe especificar una capacidad de destino para las instancias bajo demanda y las instancias de spot dentro de cada flota. Cuando se lanza el clúster, Amazon EMR aprovisiona instancias hasta que se atienden los destinos. Cuando Amazon EC2 recupera una instancia puntual de un clúster en ejecución debido a un aumento de precio o a un fallo de la instancia, Amazon EMR intenta sustituir la instancia por cualquiera de los tipos de instancias que especifique. De ese modo, resulta más sencillo recuperar la capacidad durante un pico de los precios de spot.

Puede especificar un máximo de cinco tipos de EC2 instancias de Amazon por flota para que Amazon EMR los utilice al cumplir los objetivos, o un máximo de 30 tipos de EC2 instancias de Amazon por flota cuando cree un clúster mediante la API o AWS CLI Amazon EMR y una estrategia de asignación para instancias puntuales y bajo demanda.

También puede seleccionar varias subredes para distintas zonas de disponibilidad. Cuando Amazon EMR lanza el clúster, busca en esas subredes para encontrar las instancias y las opciones de compra que especifique. Si Amazon EMR detecta un evento AWS a gran escala en una o más de las zonas de disponibilidad, Amazon EMR intenta automáticamente desviar el tráfico de las zonas de disponibilidad afectadas e intenta lanzar nuevos clústeres que usted cree en zonas de disponibilidad alternativas de acuerdo con sus selecciones. Tenga en cuenta que la selección de la zona de disponibilidad del clúster solo se realiza en el momento de crear el clúster. Los nodos del clúster existentes no se vuelven a lanzar automáticamente en una nueva zona de disponibilidad en el caso de que se produzca una interrupción en la zona de disponibilidad.

Consideraciones para trabajar con flotas de instancias

Tenga en cuenta los siguientes aspectos al utilizar flotas de instancias con Amazon EMR.

- Puede tener una flota de instancias, y solo una, por tipo de nodo (principal, básico, de tarea). Puede especificar hasta cinco tipos de EC2 instancias de Amazon para cada flota de la AWS Management Console (o un máximo de 30 tipos por flota de instancias al crear un clúster mediante la API AWS CLI o Amazon EMR y unaEstrategia de asignación para flotas de instancias).
- Amazon EMR elige uno o todos los tipos de EC2 instancias de Amazon especificados para aprovisionarlos con las opciones de compra puntual y bajo demanda.
- Establezca las capacidades de destino para instancias de spot y bajo demanda para la flota de instancias básicas y de tarea. Usa vCPU o una unidad genérica asignada a cada EC2 instancia de Amazon que cuente para los objetivos. Amazon EMR aprovisiona instancias hasta que se ha satisfecho por completo cada capacidad de destino. Para la flota principal, el destino es siempre uno.
- Puede elegir una subred (zona de disponibilidad) o un intervalo. Si elige un intervalo, Amazon EMR aprovisiona capacidad en la zona de disponibilidad que sea la mejor opción.
- Cuando se especifica una capacidad de destino para las instancias de spot:
  - Para cada tipo de instancia, especifique un precio de spot máximo. Amazon EMR aprovisiona las instancias de spot si el precio de spot está por debajo del precio de spot máximo. Solo tiene que pagar el precio de spot, que no es necesariamente el precio de spot máximo.
  - Para cada flota, defina un período de tiempo de espera para aprovisionamiento de instancias de spot. Si Amazon EMR no puede aprovisionar la capacidad de spot, puede terminar el clúster o en su lugar cambiarlo a aprovisionamiento de capacidad bajo demanda. Esto solo se aplica al aprovisionamiento de clústeres, no a su cambio de tamaño. Si el tiempo de espera finaliza durante el proceso de cambio de tamaño del clúster, las solicitudes de spot no aprovisionadas se anularán sin transferirse a la capacidad bajo demanda.
- Para cada flota, puede especificar una de las siguientes estrategias de asignación para sus instancias puntuales: optimizadas en relación precio-capacidad, optimizadas en cuanto a capacidad, precio más bajo o capacity-optimized-prioritized diversificadas en todos los grupos.
- En el caso de cada flota, puede aplicar las siguientes estrategias de asignación para sus instancias bajo demanda: la estrategia con el precio más bajo o la estrategia con prioridad.
- En el caso de cada flota con con instancias bajo demanda, puede optar por aplicar opciones de reserva de capacidad.
- Si utiliza una estrategia de asignación para flotas de instancias, se deben tener en cuenta las siguientes consideraciones al elegir subredes para su clúster de EMR:
  - Cuando Amazon EMR aprovisiona un clúster con una flota de tareas, filtra las subredes que no tienen suficientes direcciones IP disponibles para aprovisionar todas las instancias del clúster

de EMR solicitado. Esto incluye las direcciones IP necesarias para las flotas de instancias principales, básicas y de tareas durante el lanzamiento del clúster. A continuación, Amazon EMR aprovecha su estrategia de asignación para determinar el grupo de instancias, en función del tipo de instancia y las subredes restantes con direcciones IP suficientes, para lanzar el clúster.

- Si Amazon EMR no puede lanzar todo el clúster debido a que no hay suficientes direcciones IP disponibles, intentará identificar las subredes con suficientes direcciones IP libres para lanzar las flotas de instancias esenciales (principales y básicas). En estos casos, la flota de instancias de la tarea pasará a un estado suspendido, en lugar de finalizar el clúster con un error.
- Si ninguna de las subredes especificadas contiene direcciones IP suficientes para aprovisionar las flotas de instancias principales y básicas esenciales, el lanzamiento del clúster fallará y mostrará un VALIDATION_ERROR. Esto activa un evento de finalización del clúster de gravedad CRITICAL que le notifica de que el clúster no se puede lanzar. Para impedir este problema, se recomienda aumentar la cantidad de direcciones IP en las subredes.

- Si ejecuta Amazon EMR versión emr-7.7.0 o superior y utiliza una estrategia de asignación para, por ejemplo, las flotas, puede escalar el clúster hasta 4000 EC2 instancias y 14 000 volúmenes de EBS por flota de instancias. Para las versiones anteriores a la emr-7.7.0, el clúster solo se puede escalar hasta 2000 instancias y 7000 volúmenes de EBS por flota de instancias. EC2
- Al lanzar instancias bajo demanda, puede utilizar reservas de capacidad abiertas o específicas en sus cuentas principales, básicas o de tareas. Es posible que vea una capacidad insuficiente con las instancias bajo demanda cuando utiliza la estrategia de asignación para flotas de instancias. Se recomienda que especifique varios tipos de instancias para diversificar y reducir la posibilidad de que la capacidad sea insuficiente. Para obtener más información, consulte the section called “Uso de las reservas de capacidad con flotas de instancias en Amazon EMR”.

Opciones de flota de instancias

Utilice las siguientes directrices para comprender las opciones de flota de instancia.

Temas

- Configuración de las capacidades de destino
- Opciones de lanzamiento
- Opciones para varias subredes (zonas de disponibilidad)
- Configuración del nodo maestro

Configuración de las capacidades de destino

Especifique las capacidades de destino que desee para la flota de instancias secundarias y la flota de instancias de tareas. Cuando lo haga, eso determina el número de instancias bajo demanda e instancias de spot que aprovisiona Amazon EMR. Cuando se especifica una instancia, decide la cantidad de cada instancia que se tiene en cuenta para el destino. Cuando se aprovisiona una instancia bajo demanda, se tiene en cuenta para el destino bajo demanda. Lo mismo sucede para las instancias de spot. A diferencia de las flotas de instancias básicas y de tarea, la flota de instancias principales siempre es una instancia. Por tanto, la capacidad de destino para esta flota es siempre uno.

Cuando utilizas la consola, la v del tipo CPUs de EC2 instancia de Amazon se utiliza como recuento de las capacidades de destino de forma predeterminada. Puede cambiarlo a unidades genéricas y, a continuación, especificar el recuento de cada tipo de EC2 instancia. Al utilizar el AWS CLI, se asignan manualmente unidades genéricas para cada tipo de instancia.

> **Important**
>
> Cuando eliges un tipo de instancia mediante el AWS Management Console, el número de vCPU que se muestra para cada tipo de instancia es el número de núcleos virtuales de YARN para ese tipo de instancia, no el número de EC2 v CPUs para ese tipo de instancia. Para obtener más información sobre el número de v CPUs para cada tipo de instancia, consulte Tipos de EC2 instancias de Amazon.

Para cada flota, especificas hasta cinco tipos de EC2 instancias de Amazon. Si usa un clúster Estrategia de asignación para flotas de instancias y lo crea mediante la AWS CLI API de Amazon EMR, puede especificar hasta 30 tipos de EC2 instancias por flota de instancias. Amazon EMR elige cualquier combinación de estos tipos de EC2 instancias para cumplir con las capacidades objetivo. Dado que Amazon EMR desea rellenar la capacidad de destino por completo, podría producirse un sobreuso. Por ejemplo, si hay dos unidades no atendidas y Amazon EMR solo puede aprovisionar una instancia con un recuento de cinco unidades, la instancia aún se aprovisionará, lo que significa que la capacidad de destino se excederá en tres unidades.

Si reduce la capacidad de destino para cambiar el tamaño de un clúster en ejecución, Amazon EMR intenta completar tareas de aplicación y termina instancias para satisfacer el nuevo destino. Para obtener más información, consulte Terminación al completar las tareas.

Opciones de lanzamiento

En el caso de cada grupo de instancias, especifique un precio de spot máximo por cada tipo de instancia en una flota. Puede configurar este precio como un porcentaje del precio bajo demanda o como un importe específico en dólares. Amazon EMR aprovisiona instancias de spot si el precio de spot actual en una zona de disponibilidad está por debajo de su precio de spot máximo. Solo tiene que pagar el precio de spot, que no es necesariamente el precio de spot máximo.

> **Note**
>
> Las instancias de spot con una duración definida (también conocidas como bloques de spot) dejarán de estar disponibles para los nuevos clientes a partir del 1 de julio de 2021. En el caso de los clientes que hayan utilizado previamente la característica, se continuará ofreciendo soporte a las instancias de spot con una duración definida hasta el 31 de diciembre de 2022.

Disponible en Amazon EMR 5.12.1 y versiones posteriores, tiene la opción de lanzar flotas de instancias de spot y bajo demanda con una asignación de capacidad optimizada. Esta opción de estrategia de asignación se puede configurar en la antigua AWS Management Console o mediante la API`RunJobFlow`. Tenga en cuenta que no puede personalizar la estrategia de asignación en la nueva consola. El uso de la opción de estrategia de asignación requiere permisos de rol de servicio adicionales. Si utiliza el rol de servicio predeterminado de Amazon EMR y la política administrada (`EMR_DefaultRole` y `AmazonEMRServicePolicy_v2`) para el clúster, los permisos para la opción de estrategia de asignación ya están incluidos. Si no utiliza el rol de servicio ni la política administrada de Amazon EMR predeterminados, debe agregarlos para utilizar esta opción. Consulte Rol de servicio para Amazon EMR (rol de EMR).

Para obtener más información sobre las instancias puntuales, consulte Instancias puntuales en la Guía del EC2 usuario de Amazon. Para obtener más información sobre las instancias bajo demanda, consulte Instancias bajo demanda en la Guía del EC2 usuario de Amazon.

Si opta por lanzar flotas de instancias bajo demanda con la estrategia de asignación con el precio más bajo, tiene la opción de utilizar las reservas de capacidad. Las opciones de reserva de capacidad se pueden establecer mediante la API `RunJobFlow` de Amazon EMR. Las reservas de capacidad requieren permisos de rol de servicio adicionales que debe agregar para utilizar estas opciones. Consulte Permisos de estrategia de asignación. Tenga en cuenta que no puede personalizar las reservas de capacidad en la nueva consola.

Opciones para varias subredes (zonas de disponibilidad)

Cuando utilizas flotas de instancias, puedes especificar varias EC2 subredes de Amazon dentro de una VPC, cada una de las cuales corresponde a una zona de disponibilidad diferente. Si usas EC2 - Classic, especificas las zonas de disponibilidad de forma explícita. Amazon EMR identifica la mejor zona de disponibilidad para lanzar instancias de acuerdo con sus especificaciones de flota. Las instancias se aprovisionan siempre en una única zona de disponibilidad. Puede seleccionar subredes privadas o subredes públicas, pero no puede combinarlas y las subredes que especifique deben estar dentro de la misma VPC.

Configuración del nodo maestro

Dado que la flota de instancias principales es únicamente una sola instancia, su configuración es ligeramente distinta de las flotas de instancias básicas y de tarea. Solo seleccione bajo demanda o de spot para la flota de instancias principales, ya que se compone de una única instancia. Si utiliza la consola para crear la flota de instancias, la capacidad de destino para la opción de compra que seleccione se define en 1. Si usa el AWS CLI, `TargetOnDemandCapacity` defina siempre uno `TargetSpotCapacity` o en 1, según corresponda. Aún puede elegir hasta cinco tipos de instancias para la flota de instancias principal (o un máximo de 30 si utiliza la opción de estrategia de asignación para las instancias de spot o bajo demanda). Sin embargo, a diferencia de las flotas de instancias básicas y de tarea, donde Amazon EMR podría aprovisionar varias instancias de tipos distintos, Amazon EMR selecciona un único tipo de instancia para aprovisionar la flota de instancias principales.

Estrategia de asignación para flotas de instancias

Con las versiones 5.12.1 y posteriores de Amazon EMR, puede utilizar la opción de estrategia de asignación con instancias de spot y bajo demanda para cada nodo del clúster. Al crear un clúster mediante la AWS CLI API de Amazon EMR o la consola de Amazon EMR con una estrategia de asignación, puede especificar hasta 30 tipos de EC2 instancias de Amazon por flota. Con la configuración predeterminada de flota de instancias del clúster de Amazon EMR, puede tener hasta 5 tipos de instancias por flota. Se recomienda que utilice la opción de estrategia de asignación para un aprovisionamiento de clústeres más rápido, una asignación de instancias de spot más precisa y menos interrupciones de dichas instancias.

Temas

- Estrategia de asignación con instancias bajo demanda
- Estrategia de asignación con instancias de spot
- Permisos de estrategia de asignación

- Permisos de IAM necesarios para una estrategia de asignación

Estrategia de asignación con instancias bajo demanda

Las siguientes estrategias de asignación están disponibles para sus instancias bajo demanda:

`lowest-price` (predeterminado)

La estrategia de asignación de precios más bajos lanza instancias bajo demanda desde el grupo de precios más bajos que tiene capacidad disponible. Si el grupo con el precio más bajo no tiene capacidad disponible, las instancias bajo demanda provienen del siguiente grupo con el precio más bajo con capacidad disponible.

`prioritized`

La estrategia de asignación priorizada le permite especificar un valor de prioridad para cada tipo de instancia de su flota de instancias. Amazon EMR lanza las instancias bajo demanda que tienen la máxima prioridad. Si usa esta estrategia, debe configurar la prioridad para al menos un tipo de instancia. Si no configura el valor de prioridad para un tipo de instancia, Amazon EMR asigna la prioridad más baja a ese tipo de instancia. Cada flota de instancias (principal, básica o de tarea) de un clúster puede tener un valor de prioridad diferente para un tipo de instancia determinado.

> **Note**
>
> Si utiliza la estrategia de asignación capacity-optimized-prioritizedpuntual, Amazon EMR aplica las mismas prioridades tanto a las instancias bajo demanda como a las instancias puntuales cuando establece las prioridades.

Estrategia de asignación con instancias de spot

En el caso de las instancias de spot, puede elegir una de las siguientes estrategias de asignación:

`price-capacity-optimized` (recomendado)

La estrategia de asignación optimizada en función de la relación precio-capacidad lanza instancias de spot desde los grupos de instancias de spot que tienen la mayor capacidad disponible y el precio más bajo para la cantidad de instancias que se están lanzando. Como resultado, la estrategia optimizada en función de la relación precio-capacidad suele tener más

probabilidades de conseguir capacidad de Spot y ofrece tasas de interrupción más bajas. Esta es la estrategia predeterminada para la versión 6.10.0 y versiones posteriores de Amazon EMR.

`capacity-optimized`

La estrategia de asignación optimizada en función de la capacidad lanza las instancias de spot en los grupos más disponibles con la menor probabilidad de interrupción a corto plazo. Esta es una buena opción para las cargas de trabajo que podrían tener un costo de interrupción más alto asociado con el trabajo que se reinicia. Esta es la estrategia predeterminada para las versiones 6.9.0 y anteriores de Amazon EMR.

`capacity-optimized-prioritized`

La estrategia capacity-optimized-prioritized de asignación le permite especificar un valor de prioridad para cada tipo de instancia de su flota de instancias. Amazon EMR optimiza primero la capacidad, pero respeta las prioridades del tipo de instancia sobre la base del mejor esfuerzo, por ejemplo, si la prioridad no afecta significativamente la capacidad de la flota para aprovisionar una capacidad óptima. Recomendamos esta opción si tiene cargas de trabajo que deben tener una cantidad mínima de interrupciones y que aún requieren ciertos tipos de instancias. Si usa esta estrategia, debe configurar la prioridad para al menos un tipo de instancia. Si no configura una prioridad para ningún tipo de instancia, Amazon EMR asigna el valor de prioridad más bajo a ese tipo de instancia. Cada flota de instancias (principal, básica o de tarea) de un clúster puede tener un valor de prioridad diferente para un tipo de instancia determinado.

> **Note**
>
> Si utiliza la estrategia de asignación bajo demanda priorizada, Amazon EMR aplica el mismo valor de prioridad a las instancias bajo demanda y de Spot al establecer las prioridades.

`diversified`

Con la estrategia de asignación diversificada, Amazon EC2 distribuye las instancias puntuales en todos los grupos de capacidad puntual.

`lowest-price`

La estrategia de asignación de precios más bajos lanza instancias de spot desde el grupo de precios más bajos que tiene capacidad disponible. Si el grupo con el precio más bajo no tiene capacidad disponible, las instancias de spot provienen del siguiente grupo con el precio más

bajo que tenga capacidad disponible. Si un grupo se queda sin capacidad antes de cumplir con la capacidad solicitada, la EC2 flota de Amazon opta por el siguiente grupo con el precio más bajo para seguir gestionando tu solicitud. Para garantizar que se logre la capacidad deseada, es posible que reciba instancias de spot de varios grupos. Dado que esta estrategia solo tiene en cuenta el precio de la instancia y no la disponibilidad de capacidad, podría generar tasas de interrupción elevadas.

Permisos de estrategia de asignación

La opción de estrategia de asignación requiere varios permisos de IAM que se incluyen automáticamente en el rol de servicio predeterminado de Amazon EMR y en la política administrada de Amazon EMR (`EMR_DefaultRole` y `AmazonEMRServicePolicy_v2`). Si utiliza un rol de servicio o una política administrada personalizados para su clúster, debe agregar estos permisos antes de crear el clúster. Para obtener más información, consulte Permisos de estrategia de asignación.

Las reservas de capacidad bajo demanda (ODCRs) opcionales están disponibles cuando utilizas la opción de estrategia de asignación bajo demanda. Las opciones de reserva de capacidad le permiten especificar una preferencia para utilizar primero la capacidad reservada para los clústeres de Amazon EMR. Puede utilizarla para asegurarse de que sus cargas de trabajo críticas utilicen la capacidad que ya ha reservado en régimen abierto o ODCRs segmentado. En el caso de las cargas de trabajo que no son críticas, las preferencias de reserva de capacidad permiten especificar si se debe consumir la capacidad reservada.

Las reservas de capacidad solo las pueden utilizar instancias que tengan atributos coincidentes (tipo de instancia, plataforma, zona de disponibilidad). De forma predeterminada, Amazon EMR utiliza automáticamente las reservas de capacidad abiertas al aprovisionar instancias bajo demanda que coinciden con los atributos de la instancia. Si no dispone de instancias en ejecución que coincidan con los atributos de reserva de capacidad, permanecen sin utilizar hasta que lanza una instancia que coincide con sus atributos. Si no desea utilizar ninguna reserva de capacidad al lanzar el clúster, debe establecer la preferencia de reserva de capacidad como ninguna en las opciones de lanzamiento.

Sin embargo, también puede dirigir una reserva de capacidad para cargas de trabajo específicas. Esto le permite controlar de manera explícita qué instancias pueden ejecutarse en esa capacidad reservada. Para más información sobre las reservas de capacidad bajo demanda, consulte Uso de las reservas de capacidad con flotas de instancias en Amazon EMR.

Permisos de IAM necesarios para una estrategia de asignación

El Rol de servicio para Amazon EMR (rol de EMR) necesita permisos adicionales para crear un clúster que utilice la opción de estrategia de asignación para flotas de instancias de spot o bajo demanda.

Se incluyen automáticamente estos permisos en el rol de servicio predeterminado de Amazon EMR `EMR_DefaultRole` y en la política administrada de Amazon EMR `AmazonEMRServicePolicy_v2`.

Si utiliza un rol de servicio personalizado o una política administrada para su clúster, debe agregar los siguientes permisos:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DeleteLaunchTemplate",
        "ec2:CreateLaunchTemplate",
        "ec2:DescribeLaunchTemplates",
        "ec2:CreateLaunchTemplateVersion",
        "ec2:CreateFleet"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Deletelaunchtemplate"
    }
  ]
}
```

Los siguientes permisos de rol de servicio son necesarios para crear un clúster que utilice reservas de capacidad abiertas o específicas. Debe incluir estos permisos además de los permisos necesarios para utilizar la opción de estrategia de asignación.

Example Documento de política para las reservas de capacidad de los roles de servicio

Para utilizar reservas de capacidad abiertas, debe incluir los siguientes permisos adicionales.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeCapacityReservations",
        "ec2:DescribeLaunchTemplateVersions",
        "ec2:DeleteLaunchTemplateVersions"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Describecapacityreservations"
    }
  ]
}
```

Example

Para utilizar reservas de capacidad específicas, debe incluir los siguientes permisos adicionales.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeCapacityReservations",
        "ec2:DescribeLaunchTemplateVersions",
        "ec2:DeleteLaunchTemplateVersions",
        "resource-groups:ListGroupResources"
      ],
```

```
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Describecapacityreservations"
    }
  ]
}
```

Configuración de flotas de instancias para su clúster

Console

Para crear un clúster con flotas de instancias mediante la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, seleccione Crear clúster.
3. En Configuración del clúster, elija Flotas de instancias.
4. En el caso de cada grupo de nodos, seleccione Agregar tipo de instancia y elija hasta 5 tipos de instancias para las flotas de instancias principales y básicas y hasta quince tipos de instancias para las flotas de instancias de tarea. Amazon EMR podría aprovisionar cualquier combinación de estos tipos de instancia cuando se lance el clúster.
5. En cada tipo de grupo de nodos, seleccione el menú desplegable Acciones situado junto a cada instancia para cambiar estas configuraciones:

   Agregar volúmenes de EBS

   Especifique los volúmenes de EBS que desee asociar al tipo de instancia una vez que Amazon EMR lo aprovisione.

   Edición de capacidad ponderada

   En el caso del grupo de nodos principales, cambie este valor a cualquier número de unidades que se adapte a sus aplicaciones. Se utiliza el número de núcleos virtuales de YARN por cada tipo de instancias de flota como las unidades de capacidad ponderada predeterminadas. No puede editar la capacidad ponderada del nodo principal.

Edición del precio de spot máximo

Debe especificar un precio de spot máximo para cada tipo de instancia de una flota. Puede configurar este precio como un porcentaje del precio bajo demanda o como un importe específico en dólares. Si en una zona de disponibilidad el precio de spot actual está por debajo de su precio de spot máximo, Amazon EMR aprovisiona instancias de spot. Solo tiene que pagar el precio de spot, que no es necesariamente el precio de spot máximo.

6. Si lo desea, para añadir grupos de seguridad a sus nodos, expanda los grupos de EC2 seguridad (firewall) en la sección Redes y seleccione el grupo de seguridad para cada tipo de nodo.
7. De manera opcional, seleccione la casilla situada junto a Aplicar estrategia de asignación si desea utilizar la opción de estrategia de asignación y seleccione la estrategia que desee especificar para las instancias de spot. No debe seleccionar esta opción si su rol de servicio de Amazon EMR no tiene los permisos necesarios. Para obtener más información, consulte Estrategia de asignación para flotas de instancias.
8. Elija cualquier otra opción que se aplique a su clúster.
9. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para crear y lanzar un clúster con flotas de instancias con el AWS CLI, sigue estas pautas:

- Para crear y lanzar un clúster con flotas de instancias, utilice el comando `create-cluster` junto con parámetros `--instance-fleet`.
- Para obtener los detalles de configuración de las flotas de instancias en un clúster, utilice el comando `list-instance-fleets`.
- Para añadir varios Amazon Linux personalizados AMIs a un clúster que esté creando, utilice la `CustomAmiId` opción con cada `InstanceType` especificación. Puede configurar los nodos de la flota de instancias con varios tipos de instancias y varios personalizados AMIs según sus necesidades. Consulte Ejemplos: creación de un clúster con la configuración de flotas de instancia.
- Para realizar cambios en la capacidad de destino de una flota de instancias de destino, utilice el comando `modify-instance-fleet`.

- Para añadir una flota de instancias de tareas a un clúster que no dispone de una, utilice el comando `add-instance-fleet`.
- Se AMIs pueden añadir varias configuraciones personalizadas a la flota de instancias de la tarea mediante el CustomAmiId argumento del add-instance-fleet comando. Consulte Ejemplos: creación de un clúster con la configuración de flotas de instancia.
- Para utilizar la opción de estrategia de asignación al crear una flota de instancias, actualice el rol de servicio para incluir el ejemplo del documento de política en la siguiente sección.
- Para utilizar las opciones de reserva de capacidad al crear una flota de instancias con la estrategia de asignación bajo demanda, actualice el rol de servicio para incluir el ejemplo del documento de política en la siguiente sección.
- Las flotas de instancias se incluyen automáticamente en el rol de servicio de EMR predeterminada y en la política administrada de Amazon EMR (`EMR_DefaultRole` y `AmazonEMRServicePolicy_v2`). Si utiliza un rol de servicio personalizado o una política administrada personalizada para su clúster, debe agregar los nuevos permisos para la estrategia de asignación en la siguiente sección.

Ejemplos: creación de un clúster con la configuración de flotas de instancia

Los siguientes ejemplos muestran comandos `create-cluster` con una variedad de opciones que puede combinar.

> **Note**
>
> Si no ha creado anteriormente el rol de servicio Amazon EMR y el perfil de EC2 instancia predeterminados, utilícelos `aws emr create-default-roles` para crearlos antes de usar el `create-cluster` comando.

Example Ejemplo: VPC predeterminada, principal bajo demanda, básica bajo demanda con tipo de instancia único

```
aws emr create-cluster --release-label emr-5.3.1 --service-role EMR_DefaultRole \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleets \

 InstanceFleetType=MASTER,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarg
 \
```

```
 InstanceFleetType=CORE,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarge}
```

Example Ejemplo: VPC predeterminada, principal de spot, spot básica con tipo de instancia único

```
aws emr create-cluster --release-label emr-5.3.1 --service-role EMR_DefaultRole \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleets \
    InstanceFleetType=MASTER,TargetSpotCapacity=1,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5}'] \
    InstanceFleetType=CORE,TargetSpotCapacity=1,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5}']
```

Example Ejemplo: núcleo primario mixto bajo demanda con un solo tipo de instancia y una sola subred EC2

```
aws emr create-cluster --release-label emr-5.3.1 --service-role EMR_DefaultRole \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetIds=['subnet-ab12345c'] \
  --instance-fleets \
    InstanceFleetType=MASTER,TargetOnDemandCapacity=1,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge}'] \
    InstanceFleetType=CORE,TargetOnDemandCapacity=2,TargetSpotCapacity=6,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5,WeightedCapacity=2}']
```

Example Ejemplo: principal bajo demanda, núcleo puntual con varios tipos de instancias ponderadas, tiempo de espera para el spot, rango de subredes EC2

```
aws emr create-cluster --release-label emr-5.3.1 --service-role EMR_DefaultRole \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetIds=['subnet-
ab12345c','subnet-de67890f'] \
  --instance-fleets \
    InstanceFleetType=MASTER,TargetOnDemandCapacity=1,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge}'] \
    InstanceFleetType=CORE,TargetSpotCapacity=11,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5,WeightedCapacity=3}',\
'{InstanceType=m4.2xlarge,BidPrice=0.9,WeightedCapacity=5}'],\
LaunchSpecifications={SpotSpecification='{TimeoutDurationMinutes=120,TimeoutAction=SWITCH_TO_ON
```

Example Ejemplo: principal bajo demanda, de núcleo mixto y de tareas con varios tipos de instancias ponderadas, tiempo de espera para las instancias puntuales principales, rango de subredes EC2

```
aws emr create-cluster --release-label emr-5.3.1 --service-role EMR_DefaultRole \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,SubnetIds=['subnet-
ab12345c','subnet-de67890f'] \
  --instance-fleets \

 InstanceFleetType=MASTER,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=m5.xlarg
 \
    InstanceFleetType=CORE,TargetOnDemandCapacity=8,TargetSpotCapacity=6,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5,WeightedCapacity=3}',\
'{InstanceType=m4.2xlarge,BidPrice=0.9,WeightedCapacity=5}'],\
LaunchSpecifications={SpotSpecification='{TimeoutDurationMinutes=120,TimeoutAction=SWITCH_TO_ON
 \
    InstanceFleetType=TASK,TargetOnDemandCapacity=3,TargetSpotCapacity=3,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5,WeightedCapacity=3}']
```

Example Ejemplo: spot principal, sin básicos ni de tareas, configuración de Amazon EBS, VPC predeterminada

```
aws emr create-cluster --release-label Amazon EMR 5.3.1 --service-role EMR_DefaultRole
 \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleets \
    InstanceFleetType=MASTER,TargetSpotCapacity=1,\
LaunchSpecifications={SpotSpecification='{TimeoutDurationMinutes=60,TimeoutAction=TERMINATE_CLU
\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5,\
EbsConfiguration={EbsOptimized=true,EbsBlockDeviceConfigs=[{VolumeSpecification={VolumeType=gp2
\
SizeIn GB=100}},{VolumeSpecification={VolumeType=io1,SizeInGB=100,Iop
 s=100},VolumesPerInstance=4}]}}']
```

Example Ejemplo: varias instancias personalizadas AMIs, varios tipos de instancias, instancia principal bajo demanda, instancia principal bajo demanda

```
aws emr create-cluster --release-label Amazon EMR 5.3.1 --service-role EMR_DefaultRole
 \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleets \
    InstanceFleetType=MASTER,TargetOnDemandCapacity=1,\
```

```
InstanceTypeConfigs=['{InstanceType=m5.xlarge,CustomAmiId=ami-123456},
{InstanceType=m6g.xlarge, CustomAmiId=ami-234567}'] \
    InstanceFleetType=CORE,TargetOnDemandCapacity=1,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,CustomAmiId=ami-123456},
{InstanceType=m6g.xlarge, CustomAmiId=ami-234567}']
```

Example Ejemplo: añadir un nodo de tareas a un clúster en ejecución con varios tipos de instancias y varias personalizadas AMIs

```
aws emr add-instance-fleet --cluster-id j-123456 --release-label Amazon EMR 5.3.1 \
  --service-role EMR_DefaultRole \
  --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleet \
    InstanceFleetType=Task,TargetSpotCapacity=1,\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,CustomAmiId=ami-123456}',\
'{InstanceType=m6g.xlarge,CustomAmiId=ami-234567}']
```

Example Ejemplo: utilizar un archivo de configuración JSON

Puede configurar parámetros de flota de instancias en un archivo JSON y, a continuación, hacer referencia al archivo JSON como único parámetro para las flotas de instancias. Por ejemplo, el siguiente comando hace referencia a un archivo de configuración JSON, *`my-fleet-config.json`*:

```
aws emr create-cluster --release-label emr-5.30.0 --service-role EMR_DefaultRole \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole \
--instance-fleets file://my-fleet-config.json
```

El *`my-fleet-config.json`* archivo especifica las flotas de instancias principales, principales y de tareas, como se muestra en el siguiente ejemplo. La flota de instancias principales usa un precio spot máximo (`BidPrice`) como porcentaje de la demanda, mientras que las flotas de tareas y de instancia principal usan un precio spot máximo (BidPriceAsPercentageofOnDemandPrice) como cadena en USD.

```
[
    {
        "Name": "Masterfleet",
        "InstanceFleetType": "MASTER",
        "TargetSpotCapacity": 1,
        "LaunchSpecifications": {
            "SpotSpecification": {
                "TimeoutDurationMinutes": 120,
```

```
                "TimeoutAction": "SWITCH_TO_ON_DEMAND"
            }
        },
        "InstanceTypeConfigs": [
            {
                "InstanceType": "m5.xlarge",
                "BidPrice": "0.89"
            }
        ]
    },
    {
        "Name": "Corefleet",
        "InstanceFleetType": "CORE",
        "TargetSpotCapacity": 1,
        "TargetOnDemandCapacity": 1,
        "LaunchSpecifications": {
          "OnDemandSpecification": {
            "AllocationStrategy": "lowest-price",
            "CapacityReservationOptions":
            {
                "UsageStrategy": "use-capacity-reservations-first",
                "CapacityReservationResourceGroupArn": "String"
            }
        },
            "SpotSpecification": {
                "AllocationStrategy": "capacity-optimized",
                "TimeoutDurationMinutes": 120,
                "TimeoutAction": "TERMINATE_CLUSTER"
            }
        },
        "InstanceTypeConfigs": [
            {
                "InstanceType": "m5.xlarge",
                "BidPriceAsPercentageOfOnDemandPrice": 100
            }
        ]
    },
    {
        "Name": "Taskfleet",
        "InstanceFleetType": "TASK",
        "TargetSpotCapacity": 1,
        "LaunchSpecifications": {
          "OnDemandSpecification": {
            "AllocationStrategy": "lowest-price",
```

```
            "CapacityReservationOptions":
            {
                "CapacityReservationPreference": "none"
            }
        },
            "SpotSpecification": {
                "TimeoutDurationMinutes": 120,
                "TimeoutAction": "TERMINATE_CLUSTER"
            }
        },
        "InstanceTypeConfigs": [
            {
                "InstanceType": "m5.xlarge",
                "BidPrice": "0.89"
            }
        ]
    }
]
```

Modificar capacidades de destino para una flota de instancias

Utilice el comando `modify-instance-fleet` para especificar nuevas capacidades de destino para una flota de instancias. Debe especificar el ID del clúster y el ID de la flota de instancias. Usa el `list-instance-fleets` comando para recuperar la flota IDs de instancias.

```
aws emr modify-instance-fleet --cluster-id <cluster-id> \
  --instance-fleet \
    InstanceFleetId='<instance-fleet-id>',TargetOnDemandCapacity=1,TargetSpotCapacity=1
```

Agregar una flota de instancias de tarea a un clúster

Si un clúster tiene únicamente flotas de instancias principales y básicas, puede utilizar el comando `add-instance-fleet` para agregar una flota de instancias de tarea. Solo puede utilizar esto para añadir las flotas de instancias de tareas.

```
aws emr add-instance-fleet --cluster-id <cluster-id>
  --instance-fleet \
    InstanceFleetType=TASK,TargetSpotCapacity=1,\
LaunchSpecifications={SpotSpecification='{TimeoutDurationMinutes=20,TimeoutAction=TERMINATE_CLU
\
InstanceTypeConfigs=['{InstanceType=m5.xlarge,BidPrice=0.5}']
```

Obtener detalles de configuración de flotas de instancias en un clúster

Utilice el comando `list-instance-fleets` para obtener detalles de configuración de las flotas de instancias en un clúster. El comando toma un ID de clúster como entrada. El siguiente ejemplo muestra el comando y su salida de un clúster que contiene un grupo de instancias de tarea principal y un grupo de instancias de tarea básico. Para ver la sintaxis de respuesta completa, consulte ListInstanceFleetsla referencia de la API de Amazon EMR.

```
list-instance-fleets --cluster-id <cluster-id>
```

```
{
    "InstanceFleets": [
        {
            "Status": {
                "Timeline": {
                    "ReadyDateTime": 1488759094.637,
                    "CreationDateTime": 1488758719.817
                },
                "State": "RUNNING",
                "StateChangeReason": {
                    "Message": ""
                }
            },
            "ProvisionedSpotCapacity": 6,
            "Name": "CORE",
            "InstanceFleetType": "CORE",
            "LaunchSpecifications": {
                "SpotSpecification": {
                    "TimeoutDurationMinutes": 60,
                    "TimeoutAction": "TERMINATE_CLUSTER"
                }
            },
            "ProvisionedOnDemandCapacity": 2,
            "InstanceTypeSpecifications": [
                {
                    "BidPrice": "0.5",
                    "InstanceType": "m5.xlarge",
                    "WeightedCapacity": 2
                }
            ],
            "Id": "if-1ABC2DEFGHIJ3"
        },
```

```
        {
            "Status": {
                "Timeline": {
                    "ReadyDateTime": 1488759058.598,
                    "CreationDateTime": 1488758719.811
                },
                "State": "RUNNING",
                "StateChangeReason": {
                    "Message": ""
                }
            },
            "ProvisionedSpotCapacity": 0,
            "Name": "MASTER",
            "InstanceFleetType": "MASTER",
            "ProvisionedOnDemandCapacity": 1,
            "InstanceTypeSpecifications": [
                {
                    "BidPriceAsPercentageOfOnDemandPrice": 100.0,
                    "InstanceType": "m5.xlarge",
                    "WeightedCapacity": 1
                }
            ],
           "Id": "if-2ABC4DEFGHIJ4"
        }
    ]
}
```

Reconfiguración de las flotas de instancias para su clúster de Amazon EMR

Con Amazon EMR versión 5.21.0 y versiones posteriores, puede reconfigurar las aplicaciones de clúster y especificar clasificaciones de configuración adicionales para cada flota de instancias de un clúster en ejecución. Para ello, puede utilizar la interfaz de línea de AWS comandos (AWS CLI) o el SDK. AWS

Puede realizar un seguimiento del estado de una flota de instancias mediante la visualización de los CloudWatch eventos. Para obtener más información, consulte Eventos de reconfiguración de la flota de instancias.

> **Note**
>
> Solo puede anular el objeto de configuraciones del clúster especificado durante la creación del clúster. Para obtener más información sobre los objetos de configuración, consulte la sintaxis de la RunJobFlow solicitud. Si hay diferencias entre la configuración existente y el archivo que suministra, Amazon EMR restablece las configuraciones modificadas manualmente, como las que ha modificado mientras estaba conectado al clúster mediante SSH, a los valores predeterminados del clúster para la flota de instancias especificada.

Cuando envía una solicitud de reconfiguración mediante la consola de Amazon EMR, AWS la interfaz de línea de comandos AWS CLI() o el SDK, Amazon EMR comprueba AWS el archivo de configuración existente en el clúster. Si hay diferencias entre la configuración existente y el archivo que suministra, Amazon EMR inicia acciones de reconfiguración, reinicia algunas aplicaciones y restablece cualquier configuración modificada manualmente, como las configuraciones que haya modificado mientras estaba conectado al clúster mediante SSH, a los valores predeterminados del clúster para la flota de instancias especificada.

Comportamientos de reconfiguración

La reconfiguración sobrescribe la configuración del clúster con el conjunto de configuraciones recién enviado y puede sobrescribir los cambios de configuración realizados fuera de la API de reconfiguración.

Amazon EMR sigue un proceso continuo para reconfigurar las instancias de la flota de instancias principales y de tareas. Solo un porcentaje de las instancias de un único tipo de instancia se modifica y reinicia a la vez. Si su flota de instancias tiene varias configuraciones de tipos de instancias diferentes, se reconfigurarán en paralelo.

Las reconfiguraciones se declaran en el InstanceTypeConfignivel. Para ver un ejemplo visual, consulte. Reconfigure una flota de instancias Puede enviar solicitudes de reconfiguración que contengan ajustes de configuración actualizados para uno o más tipos de instancias en una sola solicitud. Debe incluir todos los tipos de instancias que forman parte de su flota de instancias en la solicitud de modificación; sin embargo, los tipos de instancia con campos de configuración rellenados se reconfigurarán, mientras que las demás `InstanceTypeConfig` instancias de la flota permanecerán sin cambios. Una reconfiguración se considera correcta solo cuando todas las instancias de los tipos de instancias especificados completan la reconfiguración. Si alguna instancia

no se reconfigura, toda la flota de instancias vuelve automáticamente a su última configuración estable conocida.

Limitaciones

Al reconfigurar una flota de instancias en un clúster en ejecución, tenga en cuenta las siguientes limitaciones:

- Las aplicaciones que no son de YARN pueden fallar durante el reinicio o provocar problemas con el clúster, especialmente si no están configuradas correctamente. Los clústeres que se acerquen al uso máximo de memoria y CPU pueden tener problemas tras el proceso de reinicio. Esto es especialmente cierto en el caso de la flota de instancias principal. Consulte la Solucione los problemas de reconfiguración de la flota de instancias sección.
- Las operaciones de redimensionamiento y reconfiguración no se realizan en paralelo. Las solicitudes de reconfiguración esperarán a que se produzca un cambio de tamaño continuo y viceversa.
- Las operaciones de redimensionamiento y reconfiguración no se realizan en paralelo. Las solicitudes de reconfiguración esperarán a que se produzca un cambio de tamaño continuo y viceversa.
- Tras volver a configurar una flota de instancias, Amazon EMR reinicia las aplicaciones para permitir que las nuevas configuraciones surtan efecto. Puede darse un error de trabajo u otro comportamiento de la aplicación inesperado si las aplicaciones están en uso durante la reconfiguración.
- Si se produce un error en la reconfiguración de cualquier tipo de instancia de una flota de instancias, Amazon EMR revierte los parámetros de configuración a la versión de trabajo anterior para toda la flota de instancias, además de emitir eventos y actualizar los detalles de estado. Si el proceso de reversión también falla, debes enviar una nueva `ModifyInstanceFleet` solicitud para recuperar la flota de instancias del estado. `ARRESTED` Los errores de reversión provocan eventos de reconfiguración de la flota de instancias y un cambio de estado.
- Las solicitudes de reconfiguración para cualquier Phoenix clasificaciones de configuración solo se admiten en la versión 5.23.0 y versiones posteriores de Amazon EMR, y no son compatibles con la versión 5.21.0 o 5.22.0 de Amazon EMR.
- Las solicitudes de reconfiguración para las clasificaciones de HBase configuración solo se admiten en la versión 5.30.0 y posteriores de Amazon EMR, y no se admiten en las versiones 5.23.0 a 5.29.0 de Amazon EMR.

- La reconfiguración de la clasificación hdfs-encryption-zones o cualquiera de las clasificaciones de configuración de Hadoop KMS no se admite en clúster de Amazon EMR con varios nodos principales.
- Amazon EMR actualmente no admite determinadas solicitudes de reconfiguración para el programador de capacidad de YARN que requieren reiniciar el YARN. ResourceManager Por ejemplo, no puede eliminar por completo una cola.
- Cuando YARN necesita reiniciarse, todos los trabajos de YARN en ejecución normalmente finalizan y se pierden. Esto puede provocar retrasos en el procesamiento de datos. Para ejecutar trabajos de YARN durante un reinicio de YARN, puedes crear un clúster de Amazon EMR con varios nodos principales o establecer yarn.resourcemanager.recovery.enabled en tu clasificación de configuración de yarn-site. `true` Para obtener más información sobre el uso de varios nodos maestros, consulte YARN de alta disponibilidad. ResourceManager

Reconfigure una flota de instancias

Using the AWS CLI

Usa el `modify-instance-fleet` comando para especificar una nueva configuración para una flota de instancias en un clúster en ejecución.

> **Note**
>
> En los siguientes ejemplos, reemplaza j-2 AL4 XXXXXX5 T9 por tu ID de clúster y reemplaza if-1xxxxxxx9 por tu ID de flota de instancias.

Ejemplo: reemplazar una configuración de una flota de instancias

> **Warning**
>
> Especifica todos `InstanceTypeConfig` los campos que usaste en el lanzamiento. Si no se incluyen los campos, es posible que se sobrescriban las especificaciones que declaraste en el lanzamiento. Consulte InstanceTypeConfigpara obtener una lista.

En el siguiente ejemplo, se hace referencia a un archivo JSON de configuración denominado InstanceFleet.json para editar la propiedad del comprobador de estado del NodeManager disco YARN para una flota de instancias.

Modificación de flota de instancias (JSON

1. Prepare la clasificación de la configuración y guárdela como InstanceFleet.json en el mismo directorio en el que ejecutará el comando.

```
{
    "InstanceFleetId":"if-1xxxxxxx9",
    "InstanceTypeConfigs": [
            {
                "InstanceType": "m5.xlarge",
               other InstanceTypeConfig fields
                "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"
                        }
                    }
                ]
            },
            {
                "InstanceType": "r5.xlarge",
               other InstanceTypeConfig fields
                "Configurations": [
                    {
                        "Classification": "yarn-site",
                        "Properties": {
                            "yarn.nodemanager.disk-health-checker.enable":"false",
                            "yarn.nodemanager.disk-health-checker.max-disk-
utilization-per-disk-percentage":"70.0"
                        }
                    }
                ]
            }
        ]
```

2. Ejecute el siguiente comando.

```
aws emr modify-instance-fleet \
--cluster-id j-2AL4XXXXXX5T9 \
--region us-west-2 \
```

```
--instance-fleet instanceFleet.json
```

Ejemplo: añadir una configuración a una flota de instancias

Si quieres añadir una configuración a un tipo de instancia, debes incluir todas las configuraciones previamente especificadas para ese tipo de instancia en la nueva `ModifyInstanceFleet` solicitud. De lo contrario, las configuraciones antes especificadas se eliminan.

En el siguiente ejemplo, se agrega una propiedad para el comprobador de memoria NodeManager virtual YARN. La configuración también incluye valores previamente especificados para el comprobador de estado del NodeManager disco YARN, de modo que los valores no se sobrescriban.

1. Prepara el siguiente contenido en InstanceFleet.json y guárdalo en el mismo directorio en el que ejecutarás el comando.

```
{
    "InstanceFleetId":"if-1xxxxxxx9",
    "InstanceTypeConfigs": [
            {
                "InstanceType": "m5.xlarge",
                other InstanceTypeConfig fields
                "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"
                        }
                    }
                ]
            },
            {
                "InstanceType": "r5.xlarge",
                other InstanceTypeConfig fields
                "Configurations": [
                    {
                        "Classification": "yarn-site",
```

```
                        "Properties": {
                            "yarn.nodemanager.disk-health-checker.enable":"false",
                            "yarn.nodemanager.disk-health-checker.max-disk-
utilization-per-disk-percentage":"70.0"
                        }
                    }
                ]
            }
        ]
}
```

2. Ejecute el siguiente comando.

```
aws emr modify-instance-fleet \
--cluster-id j-2AL4XXXXXX5T9 \
--region us-west-2 \
--instance-fleet instanceFleet.json
```

using the Java SDK

> **Note**
>
> En los siguientes ejemplos, reemplaza j-2 AL4 XXXXXX5 T9 por tu ID de clúster y reemplaza if-1xxxxxxx9 por tu ID de flota de instancias.

El siguiente fragmento de código proporciona una nueva configuración para una flota de instancias mediante el AWS SDK for Java.

```
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 newConfiguration = new Configuration()
    .withClassification("hive-site")
    .withProperties(hiveProperties);

List<InstanceTypeConfig> instanceTypeConfigList = new ArrayList<>();
```

```
for (InstanceTypeConfig instanceTypeConfig : currentInstanceTypeConfigList) {
    instanceTypeConfigList.add(new InstanceTypeConfig()
        .withInstanceType(instanceTypeConfig.getInstanceType())
        .withBidPrice(instanceTypeConfig.getBidPrice())
        .withWeightedCapacity(instanceTypeConfig.getWeightedCapacity())
        .withConfigurations(newConfiguration)
    );
}

InstanceFleetModifyConfig instanceFleetModifyConfig = new
 InstanceFleetModifyConfig()
    .withInstanceFleetId("if-1xxxxxxx9")
    .withInstanceTypeConfigs(instanceTypeConfigList);

ModifyInstanceFleetRequest modifyInstanceFleetRequest = new
 ModifyInstanceFleetRequest()
    .withInstanceFleet(instanceFleetModifyConfig)
    .withClusterId("j-2AL4XXXXXX5T9");

emrClient.modifyInstanceFleet(modifyInstanceFleetRequest);
```

Solucione los problemas de reconfiguración de la flota de instancias

Si se produce un error en el proceso de reconfiguración de cualquier tipo de instancia de una flota de instancias, Amazon EMR revierte la reconfiguración en curso y registra un mensaje de error mediante un evento de eventos. AAmazon CloudWatch El evento proporciona un breve resumen del error de reconfiguración. Enumera las instancias en las que la reconfiguración ha fallado y los mensajes de error correspondientes. A continuación se muestra el ejemplo de un mensaje de error.

```
Amazon EMR couldn't revert the instance fleet if-1xxxxxxx9 in the Amazon
EMR cluster j-2AL4XXXXXX5T9 (ExampleClusterName) to the previously
successful configuration at 2021-01-01 00:00 UTC. The reconfiguration
reversion failed because of Instance i-xxxxxxx1, i-xxxxxxx2, i-xxxxxxx3
failed with message "This is an example failure message"...
```

Para acceder a los registros de aprovisionamiento de nodos

Use SSH para conectarse al nodo en el que no se pudo realizar la reconfiguración. Para obtener instrucciones, consulte Connect to your Linux instance in Amazon Elastic Compute Cloud.

Accessing logs by connecting to a node

1. Navegue hasta el siguiente directorio, que contiene los archivos de registro de aprovisionamiento del nodo.

```
/mnt/var/log/provision-node/
```

2. Abre el subdirectorio de informes y busca el informe de aprovisionamiento de nodos para la reconfiguración. El directorio de informes organiza los registros por número de versión de reconfiguración, identificador único universal (UUID), dirección IP de la EC2 instancia de Amazon y marca de tiempo. Cada informe es un archivo YAML comprimido que contiene información detallada sobre el proceso de reconfiguración. A continuación, se muestra un ejemplo del nombre y la ruta de un archivo de informe.

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
xxx.ec2.internal/202104061715.yaml.gz
```

3. Puede examinar un informe con un visor de archivos como zless, como en el siguiente ejemplo.

```
zless 202104061715.yaml.gz
```

Accessing logs using Amazon S3

Inicie sesión en la consola de Amazon S3 AWS Management Console y ábrala en https://console.aws.amazon.com/s3/. Abra el bucket de Amazon S3 que ha especificado al configurar el clúster para archivar los archivos de registro.

1. Navegue hasta la siguiente carpeta, que contiene los archivos de registro de aprovisionamiento del nodo:

```
amzn-s3-demo-bucket/elasticmapreduce/cluster id/node/instance id/provision-node/
```

2. Abra la carpeta de informes y busque el informe de aprovisionamiento de nodos para la reconfiguración. La carpeta de informes organiza los registros por número de versión de reconfiguración, identificador único universal (UUID), dirección IP de la EC2 instancia de Amazon y marca de tiempo. Cada informe es un archivo YAML comprimido que contiene información detallada sobre el proceso de reconfiguración. A continuación, se muestra un ejemplo del nombre y la ruta de un archivo de informe.

```
/reports/2/ca598xxx-cxxx-4xxx-bxxx-6dbxxxxxxxxx/ip-10-73-xxx-
xxx.ec2.internal/202104061715.yaml.gz
```

Para ver un archivo de registro, puede descargarlo de Amazon S3 en su máquina local como un archivo de texto. Para obtener instrucciones, consulte Descarga de un objeto.

Cada archivo de registro contiene un informe de aprovisionamiento detallado para la reconfiguración asociada. Para encontrar información sobre los mensajes de error, puede buscar el nivel de registro `err` de un informe. El formato del informe depende de la versión de Amazon EMR del clúster. El siguiente ejemplo muestra información de error para las versiones 5.32.0 y 6.2.0 de Amazon EMR y las versiones posteriores utilizan el siguiente formato:

```
- level: err
  message: 'Example detailed error message.'
  source: Puppet
  tags:
  - err
  time: '2021-01-01 00:00:00.000000 +00:00'
  file:
  line:
```

Uso de las reservas de capacidad con flotas de instancias en Amazon EMR

Para lanzar flotas de instancias bajo demanda con opciones de reserva de capacidad, asocie los permisos de rol de servicio adicionales necesarios para utilizar las opciones de reserva de capacidad. Dado que las opciones de reserva de capacidad deben utilizarse junto con la estrategia de asignación bajo demanda, también debe incluir los permisos necesarios para la estrategia de asignación en su rol de servicio y en su política administrada. Para obtener más información, consulte Permisos de estrategia de asignación.

Amazon EMR admite reservas de capacidad abiertas y específicas. En los temas siguientes se muestran las configuraciones de las flotas de instancias que puede utilizar con la acción `RunJobFlow` o el comando `create-cluster` para lanzar flotas de instancias mediante reservas de capacidad bajo demanda.

Uso de reservas de capacidad abiertas en la medida de lo posible

Si las instancias bajo demanda del clúster coinciden con los atributos de las reservas de capacidad abiertas (tipo de instancia, plataforma, tenencia y zona de disponibilidad) disponibles en su cuenta, las reservas de capacidad se aplican automáticamente. Sin embargo, no se garantiza que se utilicen sus reservas de capacidad. Para aprovisionar el clúster, Amazon EMR evalúa todos los grupos de instancias especificados en la solicitud de lanzamiento y utiliza el que tenga el precio más bajo y la capacidad suficiente para lanzar todos los nodos principales solicitados. Las reservas de capacidad abiertas disponibles que coincidan con el grupo de instancias se aplican automáticamente. Si las reservas de capacidad abiertas disponibles no coinciden con el grupo de instancias, no se utilizarán.

Una vez aprovisionados los nodos principales, se selecciona y se fija la zona de disponibilidad. Amazon EMR aprovisiona los nodos de tarea en grupos de instancias, empezando por los de menor precio, en la zona de disponibilidad seleccionada hasta que se aprovisionen todos los nodos de este tipo. Las reservas de capacidad abiertas disponibles que coincidan con los grupos de instancias se aplican automáticamente.

A continuación se incluyen casos de uso de la lógica de asignación de capacidad de Amazon EMR para utilizar reservas de capacidad abiertas en la medida de lo posible.

Ejemplo 1: el grupo de instancias con el precio más bajo de la solicitud de lanzamiento tiene reservas de capacidad abiertas disponibles

En este caso, Amazon EMR lanza la capacidad en el grupo de instancias de menor precio con instancias bajo demanda. Las reservas de capacidad abiertas disponibles del grupo de instancias se usan automáticamente.

Estrategia bajo demanda	lowest-price		
Capacidad solicitada	100		
Tipo de instancia	c5.xlarge	m5.xlarge	r5.xlarge
Reservas de capacidad abiertas disponibles	150	100	100
Precio bajo demanda	$	$$	$$$

Instancias aprovisionadas	100	-	-
Reserva de capacidad abierta utilizada	100	-	-
Reservas de capacidad abiertas disponibles	50	100	100

Después de lanzar la flota de instancias, puede ejecutar `describe-capacity-reservations` para ver cuántas reservas de capacidad sin utilizar quedan.

Ejemplo 2: el grupo de instancias con el precio más bajo de la solicitud de lanzamiento no tiene reservas de capacidad abiertas disponibles

En este caso, Amazon EMR lanza la capacidad en el grupo de instancias de menor precio con instancias bajo demanda. Sin embargo, sus reservas de capacidad abiertas no se utilizan.

Estrategia bajo demanda	lowest-price		
Capacidad solicitada	100		
Tipo de instancia	c5.xlarge	m5.xlarge	r5.xlarge
Reservas de capacidad abiertas disponibles	-	-	100
Precio bajo demanda	$	$$	$$$
Instancias aprovisionadas	100	-	-
Reserva de capacidad abierta utilizada	-	-	-

Reservas de capacidad abiertas disponibles	-	-	100

Configuración de flotas de instancias para utilizar las reservas de capacidad abiertas en la medida de lo posible

Cuando utilice la acción `RunJobFlow` para crear un clúster basado en una flota de instancias, defina la estrategia de asignación bajo demanda en `lowest-price` y `CapacityReservationPreference` para las opciones de reserva de capacidad en `open`. Como alternativa, si deja este campo en blanco, Amazon EMR establece de forma predeterminada la preferencia de reserva de capacidad de la instancia bajo demanda en `open`.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "CapacityReservationPreference": "open"
         }
        }
    }
```

También puede usar la CLI de Amazon EMR para crear un clúster basado en una flota de instancias mediante reservas de capacidad abiertas.

```
aws emr create-cluster \
 --name 'open-ODCR-cluster' \
 --release-label emr-5.30.0 \
 --service-role EMR_DefaultRole \
 --ec2-attributes SubnetId=subnet-22XXXX01,InstanceProfile=EMR_EC2_DefaultRole \
 --instance-fleets
 InstanceFleetType=MASTER,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=c4.xlarg
 \

 InstanceFleetType=CORE,TargetOnDemandCapacity=100,InstanceTypeConfigs=['{InstanceType=c5.xlarg
{InstanceType=m5.xlarge},{InstanceType=r5.xlarge}'],\
   LaunchSpecifications={OnDemandSpecification='{AllocationStrategy=lowest-
price,CapacityReservationOptions={CapacityReservationPreference=open}}'}
```

Donde,

- `open-ODCR-cluster` se sustituye por el nombre del clúster que utiliza reservas de capacidad abiertas.
- `subnet-22XXXX01` se sustituye por el ID de la subred.

Uso de las reservas de capacidad abiertas en primer lugar

Puede optar por anular la estrategia de asignación de precios más bajos y priorizar el uso de las reservas de capacidad abiertas disponibles al aprovisionar un clúster de Amazon EMR. En este caso, Amazon EMR evalúa todos los grupos de instancias con reservas de capacidad especificadas en la solicitud de lanzamiento y utiliza el que tenga el precio más bajo y la capacidad suficiente para lanzar todos los nodos básicos solicitados. Si ninguno de los grupos de instancias con reservas de capacidad tiene capacidad suficiente para los nodos principales solicitados, Amazon EMR recurrirá al caso en el que se hace todo lo posible que se ha descrito en el tema anterior. Es decir, Amazon EMR reevalúa todos los grupos de instancias especificados en la solicitud de lanzamiento y utiliza el que tenga el precio más bajo y la capacidad suficiente para lanzar todos los nodos básicos solicitados. Las reservas de capacidad abiertas disponibles que coincidan con el grupo de instancias se aplican automáticamente. Si las reservas de capacidad abiertas disponibles no coinciden con el grupo de instancias, no se utilizarán.

Una vez aprovisionados los nodos principales, se selecciona y se fija la zona de disponibilidad. Amazon EMR aprovisiona los nodos de tarea en grupos de instancias con reservas de capacidad, empezando por los de menor precio, en la zona de disponibilidad seleccionada hasta que se aprovisionen todos los nodos de este tipo. Amazon EMR utiliza primero las reservas de capacidad abiertas disponibles en cada grupo de instancias de la zona de disponibilidad seleccionada y, solo si es necesario, utiliza la estrategia de menor precio para aprovisionar los nodos de tarea restantes.

A continuación se incluyen casos de uso de la lógica de asignación de capacidad de Amazon EMR para utilizar primero las reservas de capacidad abiertas.

Ejemplo 1: el grupo de instancias con reservas de capacidad abiertas disponibles en la solicitud de lanzamiento tiene capacidad suficiente para los nodos principales

En este caso, Amazon EMR lanza la capacidad en el grupo de instancias con las reservas de capacidad abiertas disponibles, independientemente del precio del grupo de instancias. Como resultado, las reservas de capacidad abiertas se utilizan siempre que es posible, hasta que se aprovisionen todos los nodos principales.

Estrategia bajo demanda	lowest-price		
Capacidad solicitada	100		
Estrategia de uso	use-capacity-reservations-first		
Tipo de instancia	c5.xlarge	m5.xlarge	r5.xlarge
Reservas de capacidad abiertas disponibles	-	-	150
Precio bajo demanda	$	$$	$$$
Instancias aprovisionadas	-	-	100
Reserva de capacidad abierta utilizada	-	-	100
Reservas de capacidad abiertas disponibles	-	-	50

Ejemplo 2: el grupo de instancias con reservas de capacidad abiertas disponibles en la solicitud de lanzamiento no tiene capacidad suficiente para los nodos básicos

En este caso, Amazon EMR recurre al lanzamiento de nodos principales con la estrategia de menor precio y haciendo todo lo posible por utilizar las reservas de capacidad.

Estrategia bajo demanda	lowest-price		
Capacidad solicitada	100		
Estrategia de uso	use-capacity-reservations-first		
Tipo de instancia	c5.xlarge	m5.xlarge	r5.xlarge

Reservas de capacidad abiertas disponibles	10	50	50
Precio bajo demanda	$	$$	$$$
Instancias aprovisionadas	100	-	-
Reserva de capacidad abierta utilizada	10	-	-
Reservas de capacidad abiertas disponibles	-	50	50

Después de lanzar la flota de instancias, puede ejecutar `describe-capacity-reservations` para ver cuántas reservas de capacidad sin utilizar quedan.

Configuración de las flotas de instancias para que utilicen las reservas de capacidad abiertas en primer lugar

Cuando utilice la acción `RunJobFlow` para crear un clúster basado en una flota de instancias, defina la estrategia de asignación bajo demanda en `lowest-price` y `UsageStrategy` para `CapacityReservationOptions` en `use-capacity-reservations-first`.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "UsageStrategy": "use-capacity-reservations-first"
         }
       }
    }
```

También puede utilizar la CLI de Amazon EMR para crear un clúster basado en una flota de instancias mediante reservas de capacidad abiertas.

```
aws emr create-cluster \
```

```
  --name 'use-CR-first-cluster' \
  --release-label emr-5.30.0 \
  --service-role EMR_DefaultRole \
  --ec2-attributes SubnetId=subnet-22XXXX01,InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleets \

 InstanceFleetType=MASTER,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=c4.xlarg
 \

 InstanceFleetType=CORE,TargetOnDemandCapacity=100,InstanceTypeConfigs=['{InstanceType=c5.xlarg
{InstanceType=m5.xlarge},{InstanceType=r5.xlarge}'],\
LaunchSpecifications={OnDemandSpecification='{AllocationStrategy=lowest-
price,CapacityReservationOptions={UsageStrategy=use-capacity-reservations-first}}'}
```

Donde,

- `use-CR-first-cluster` se sustituye por el nombre del clúster que utiliza reservas de capacidad abiertas.
- `subnet-22XXXX01` se sustituye por el ID de la subred.

Uso de reservas de capacidad específicas en primer lugar

Al aprovisionar un clúster de Amazon EMR, puede optar por anular la estrategia de asignación de precios más bajos y priorizar el uso de las reservas de capacidad específicas disponibles. En este caso, Amazon EMR evalúa todos los grupos de instancias con reservas de capacidad específicas en la solicitud de lanzamiento y elige el que tenga el precio más bajo y la capacidad suficiente para lanzar todos los nodos básicos solicitados. Si ninguno de los grupos de instancias con reservas de capacidad específicas tiene capacidad suficiente para los nodos principales, Amazon EMR recurrirá al caso en el que se hace todo lo posible que se ha descrito previamente. Es decir, Amazon EMR reevalúa todos los grupos de instancias especificados en la solicitud de lanzamiento y selecciona el que tenga el precio más bajo y la capacidad suficiente para lanzar todos los nodos básicos solicitados. Las reservas de capacidad abiertas disponibles que coincidan con el grupo de instancias se aplican automáticamente. Sin embargo, sus reservas de capacidad específicas no se utilizan.

Una vez aprovisionados los nodos principales, se selecciona y se fija la zona de disponibilidad. Amazon EMR aprovisiona los nodos de tarea en grupos de instancias con reservas de capacidad específicas, empezando por los de menor precio, en la zona de disponibilidad seleccionada hasta que se aprovisionen todos los nodos de este tipo. Amazon EMR intenta utilizar primero las reservas de capacidad específicas disponibles en cada grupo de instancias de la zona de disponibilidad

seleccionada. Luego, solo si es necesario, Amazon EMR utiliza la estrategia de menor precio para aprovisionar los nodos de tarea restantes.

A continuación se indican casos de uso de la lógica de asignación de capacidad de Amazon EMR para utilizar primero las reservas de capacidad específicas.

Ejemplo 1: el grupo de instancias con reservas de capacidad específicas disponibles en la solicitud de lanzamiento tiene capacidad suficiente para los nodos principales

En este caso, Amazon EMR lanza la capacidad en el grupo de instancias con las reservas de capacidad específicas disponibles, independientemente del precio del grupo de instancias. Como resultado, las reservas de capacidad específicas se utilizan siempre que es posible, hasta que se aprovisionen todos los nodos básicos.

Estrategia bajo demanda	lowest-price		
Estrategia de uso	use-capacity-reservations-first		
Capacidad solicitada	100		
Tipo de instancia	c5.xlarge	m5.xlarge	r5.xlarge
Reservas de capacidad específicas disponibles	-	-	150
Precio bajo demanda	$	$$	$$$
Instancias aprovisionadas	-	-	100
Reserva de capacidad específica utilizada	-	-	100
Reservas de capacidad específicas disponibles	-	-	50

Example Ejemplo 2: el grupo de instancias con reservas de capacidad específicas disponibles en la solicitud de lanzamiento no tiene capacidad suficiente para los nodos básicos

Estrategia bajo demanda	lowest-price		
Capacidad solicitada	100		
Estrategia de uso	use-capacity-reservations-first		
Tipo de instancia	c5.xlarge	m5.xlarge	r5.xlarge
Reservas de capacidad específicas disponibles	10	50	50
Precio bajo demanda	$	$$	$$$
Instancias aprovisionadas	100	-	-
Reservas de capacidad específicas utilizadas	10	-	-
Reservas de capacidad específicas disponibles	-	50	50

Después de lanzar la flota de instancias, puede ejecutar `describe-capacity-reservations` para ver cuántas reservas de capacidad sin utilizar quedan.

Configuración de flotas de instancias para que utilicen primero las reservas de capacidad específicas

Cuando utilice la acción `RunJobFlow` para crear un clúster basado en una flota de instancias, defina la estrategia de asignación bajo demanda en `lowest-price`, `UsageStrategy` para `CapacityReservationOptions` en `use-capacity-reservations-first` y `CapacityReservationResourceGroupArn` para `CapacityReservationOptions` en `<your`

`resource group ARN>`. Para obtener más información, consulta Cómo trabajar con reservas de capacidad en la Guía del EC2 usuario de Amazon.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "UsageStrategy": "use-capacity-reservations-first",
            "CapacityReservationResourceGroupArn": "arn:aws:resource-groups:sa-
east-1:123456789012:group/MyCRGroup"
         }
       }
    }
```

Donde `arn:aws:resource-groups:sa-east-1:123456789012:group/MyCRGroup` se sustituye por el ARN del grupo de recursos.

También puede utilizar la CLI de Amazon EMR para crear un clúster basado en una flota de instancias mediante reservas de capacidad específicas.

```
aws emr create-cluster \
  --name 'targeted-CR-cluster' \
  --release-label emr-5.30.0 \
  --service-role EMR_DefaultRole \
  --ec2-attributes SubnetId=subnet-22XXXX01,InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleets
 InstanceFleetType=MASTER,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=c4.xlarg
 \
    InstanceFleetType=CORE,TargetOnDemandCapacity=100,\
InstanceTypeConfigs=['{InstanceType=c5.xlarge},{InstanceType=m5.xlarge},
{InstanceType=r5.xlarge}'],\
LaunchSpecifications={OnDemandSpecification='{AllocationStrategy=lowest-
price,CapacityReservationOptions={UsageStrategy=use-capacity-reservations-
first,CapacityReservationResourceGroupArn=arn:aws:resource-groups:sa-
east-1:123456789012:group/MyCRGroup}}'}
```

Donde,

- `targeted-CR-cluster` se sustituye por el nombre del clúster que utiliza reservas de capacidad específicas.
- `subnet-22XXXX01` se sustituye por el ID de la subred.

- `arn:aws:resource-groups:sa-east-1:123456789012:group/MyCRGroup` se sustituye por el ARN del grupo de recursos.

Evitar el uso de reservas de capacidad abiertas disponibles

Example

Si quiere evitar utilizar inesperadamente alguna de sus reservas de capacidad abiertas al lanzar un clúster de Amazon EMR, defina la estrategia de asignación bajo demanda en `lowest-price` y `CapacityReservationPreference` para `CapacityReservationOptions` en `none`. De lo contrario, Amazon EMR establece de forma predeterminada la preferencia de reserva de capacidad de la instancia bajo demanda en `open` e intenta utilizar las reservas de capacidad abiertas disponibles en la medida de lo posible.

```
"LaunchSpecifications":
    {"OnDemandSpecification": {
        "AllocationStrategy": "lowest-price",
        "CapacityReservationOptions":
         {
            "CapacityReservationPreference": "none"
         }
       }
    }
```

También puede usar la CLI de Amazon EMR para crear un clúster basado en una flota de instancias sin utilizar reservas de capacidad abiertas.

```
aws emr create-cluster \
  --name 'none-CR-cluster' \
  --release-label emr-5.30.0 \
  --service-role EMR_DefaultRole \
  --ec2-attributes SubnetId=subnet-22XXXX01,InstanceProfile=EMR_EC2_DefaultRole \
  --instance-fleets \

 InstanceFleetType=MASTER,TargetOnDemandCapacity=1,InstanceTypeConfigs=['{InstanceType=c4.xlarg
 \

 InstanceFleetType=CORE,TargetOnDemandCapacity=100,InstanceTypeConfigs=['{InstanceType=c5.xlarg
{InstanceType=m5.xlarge},{InstanceType=r5.xlarge}'],\
LaunchSpecifications={OnDemandSpecification='{AllocationStrategy=lowest-
price,CapacityReservationOptions={CapacityReservationPreference=none}}'}
```

Donde,

- `none-CR-cluster` se sustituye por el nombre del clúster que no utiliza reservas de capacidad abiertas.
- `subnet-22XXXX01` se sustituye por el ID de la subred.

Escenarios de uso de reservas de capacidad

Puede beneficiarse del uso de reservas de capacidad en los siguientes escenarios.

Escenario 1: rotar un clúster de ejecución prolongada con reservas de capacidad

Al rotar un clúster de ejecución prolongada, es posible que tenga requisitos estrictos en cuanto a los tipos de instancias y las zonas de disponibilidad para las nuevas instancias que aprovisione. Con las reservas de capacidad, puede utilizar la garantía de capacidad para completar la rotación del clúster sin interrupciones.

Escenario 2: aprovisionar clústeres sucesivos de corta duración con reservas de capacidad

También puede utilizar las reservas de capacidad para aprovisionar un grupo de clústeres sucesivos de corta duración para cargas de trabajo individuales, de modo que, cuando termine un clúster, el siguiente pueda utilizar las reservas de capacidad. Puede utilizar reservas de capacidad específicas para garantizar que solo los clústeres previstos utilicen las reservas de capacidad.

Configuración de grupos de instancias uniformes para su clúster de Amazon EMR

Con la configuración de grupos de instancias, cada tipo de nodo (principal, secundario o de tareas) se compone del mismo tipo de instancia y las mismas opciones de compra para las instancias: bajo demanda o de spot. Estos ajustes se especifican al crear un grupo de instancias. No se pueden cambiar más adelante. Puede, sin embargo, añadir instancias del mismo tipo y opción de compra a los grupos instancias secundarias y de tareas. También puede eliminar instancias.

Si las instancias bajo demanda del clúster coinciden con los atributos de las reservas de capacidad abiertas (tipo de instancia, plataforma, tenencia y zona de disponibilidad) disponibles en su cuenta, las reservas de capacidad se aplican automáticamente. Puede usar reservas de capacidad abiertas para los nodos principales, básicos y de tarea. Sin embargo, no puede utilizar reservas de capacidad específicas ni impedir que las instancias se lancen en reservas de capacidad abiertas con atributos

coincidentes al aprovisionar clústeres mediante grupos de instancias. Si quiere utilizar reservas de capacidad específicas o evitar que las instancias se lancen en reservas de capacidad abiertas, utilice flotas de instancias en su lugar. Para obtener más información, consulte Uso de las reservas de capacidad con flotas de instancias en Amazon EMR.

Para agregar distintos tipos de instancia una vez creado un clúster, puede agregar grupos de instancia de tareas adicionales. Puede elegir distintos tipos de instancia y opciones de compra para cada grupo de instancias. Para obtener más información, consulte Utilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes.

Al lanzar instancias, la preferencia de reserva de capacidad de una instancia bajo demanda se configura de manera predeterminada en `open`, lo que le permite ejecutarse en cualquier reserva de capacidad abierta que tenga atributos coincidentes (tipo de instancia, plataforma, zona de disponibilidad). Para más información sobre las reservas de capacidad bajo demanda, consulte Uso de las reservas de capacidad con flotas de instancias en Amazon EMR.

En esta sección se explica cómo crear un clúster con grupos de instancias uniformes. Para obtener más información sobre la modificación de un grupo de instancias existente añadiendo o eliminando instancias de forma manual o con escalado automático, consulte Administración de clústeres de Amazon EMR.

Usar la consola para configurar grupos de instancias uniformes

Console

Para crear un clúster con grupos de instancias mediante la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, seleccione Crear clúster.
3. En Configuración del clúster, elija Grupos de instancias.
4. En Grupos de nodos, hay una sección para cada tipo de grupo de nodos. Para el grupo de nodos principales, seleccione la casilla Usar varios nodos principales si desea tener 3 nodos principales. Seleccione la casilla Usar la opción de compra de spot si desea utilizar la opción de compra de spot.
5. Para los grupos de nodos principales y básicos, seleccione Agregar tipo de instancia y elija hasta 5 tipos de instancias. Para el grupo de tarea, seleccione Agregar tipo de

instancia y elija hasta quince tipos de instancias. Amazon EMR podría aprovisionar cualquier combinación de estos tipos de instancia cuando se lance el clúster.

6. En cada tipo de grupo de nodos, seleccione el menú desplegable Acciones situado junto a cada instancia para cambiar estas configuraciones:

   Agregar volúmenes de EBS

   Especifique los volúmenes de EBS que desee asociar al tipo de instancia una vez que Amazon EMR lo aprovisione.

   Edición del precio de spot máximo

   Debe especificar un precio de spot máximo para cada tipo de instancia de una flota. Puede configurar este precio como un porcentaje del precio bajo demanda o como un importe específico en dólares. Si en una zona de disponibilidad el precio de spot actual está por debajo de su precio de spot máximo, Amazon EMR aprovisiona instancias de spot. Solo tiene que pagar el precio de spot, que no es necesariamente el precio de spot máximo.
7. De manera opcional, expanda Configuración de nodos para introducir una configuración de JSON o para cargar JSON desde Amazon S3.
8. Elija cualquier otra opción que se aplique a su clúster.
9. Para lanzar el clúster, elija Crear clúster.

Úsala AWS CLI para crear un clúster con grupos de instancias uniformes

Para especificar la configuración de grupos de instancias para un clúster con la AWS CLI, utilice el comando `create-cluster` junto con el parámetro `--instance-groups`. Amazon EMR supone la opción de instancia bajo demanda, a menos que especifique el argumento `BidPrice` para un grupo de instancias. Para obtener ejemplos de comandos `create-cluster` que lanzan grupos de instancias uniformes con instancias bajo demanda y diversas opciones de clúster, escriba `aws emr create-cluster help` en la línea de comando o consulte create-cluster en la referencia de comandos de la AWS CLI .

Puede usarlo AWS CLI para crear grupos de instancias uniformes en un clúster que utilice instancias puntuales. El precio de spot que se ofrece depende de la zona de disponibilidad. Cuando utiliza la CLI o la API, puede especificar la zona de disponibilidad con el `AvailabilityZone` argumento (si utiliza una red EC2 clásica) o con el `SubnetID` argumento del `--ec2-attributes` parámetro. La zona de disponibilidad o la subred que seleccione se aplica al clúster, por lo que se utiliza para

todos los grupos de instancias. Si no especifica ninguna zona de disponibilidad o subred de forma explícita, Amazon EMR selecciona la zona de disponibilidad con el menor precio de spot al lanzar el clúster.

El siguiente ejemplo muestra un comando `create-cluster` que crea grupos de instancias principales, básicas y de tarea que utilizan instancias de spot. *myKey*Sustitúyelo por el nombre de tu EC2 key pair de Amazon.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "MySpotCluster" \
  --release-label emr-7.10.0 \
  --use-default-roles \
  --ec2-attributes KeyName=myKey \
  --instance-groups \
    InstanceGroupType=MASTER,InstanceType=m5.xlarge,InstanceCount=1,BidPrice=0.25 \
    InstanceGroupType=CORE,InstanceType=m5.xlarge,InstanceCount=2,BidPrice=0.03 \
    InstanceGroupType=TASK,InstanceType=m5.xlarge,InstanceCount=4,BidPrice=0.03 \
    InstanceGroupType=TASK,InstanceType=m5.xlarge,InstanceCount=2,BidPrice=0.04
```

Con la CLI, puede crear clústeres de grupos de instancias uniformes que especifiquen una AMI personalizada única para cada tipo de instancia del grupo de instancias. Esto le permite utilizar diferentes arquitecturas de instancias en el mismo grupo de instancias. Cada tipo de instancia debe usar una AMI personalizada con una arquitectura coincidente. Por ejemplo, se configuraría un tipo de instancia m5.xlarge con una AMI personalizada de arquitectura x86_64 y un tipo de instancia m6g.xlarge con la AMI personalizada de arquitectura `AWS AARCH64` correspondiente (ARM).

En el siguiente ejemplo, se muestra un clúster de grupos de instancias uniforme creado con dos tipos de instancias, cada uno con su propia AMI personalizada. Ten en cuenta que las AMIs personalizaciones se especifican solo a nivel de tipo de instancia, no a nivel de clúster. Esto es para evitar conflictos entre el tipo de instancia AMIs y una AMI a nivel de clúster, lo que provocaría un error en el lanzamiento del clúster.

```
aws emr create-cluster
```

```
  --release-label emr-5.30.0 \
  --service-role EMR_DefaultRole \
  --ec2-attributes SubnetId=subnet-22XXXX01,InstanceProfile=EMR_EC2_DefaultRole \
  --instance-groups \

 InstanceGroupType=MASTER,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-123456
 \

 InstanceGroupType=CORE,InstanceType=m6g.xlarge,InstanceCount=1,CustomAmiId=ami-234567
```

Puedes añadir varias instancias personalizadas AMIs a un grupo de instancias y añadirlas a un clúster en ejecución. El argumento `CustomAmiId` se puede usar con el comando `add-instance-groups`, como se muestra en el siguiente ejemplo.

```
aws emr add-instance-groups --cluster-id j-123456 \
  --instance-groups \

 InstanceGroupType=Task,InstanceType=m5.xlarge,InstanceCount=1,CustomAmiId=ami-123456
```

Uso del SDK de Java para crear un grupo de instancias

Puede instanciar un objeto `InstanceGroupConfig` que especifica la configuración de un grupo de instancias para un clúster. Para utilizar instancias de spot, define las propiedades `withBidPrice` y `withMarket` en el objeto `InstanceGroupConfig`. El código siguiente muestra cómo definir grupos de instancias principales, básicas y de tarea que ejecutan instancias de spot.

```
InstanceGroupConfig instanceGroupConfigMaster = new InstanceGroupConfig()
 .withInstanceCount(1)
 .withInstanceRole("MASTER")
 .withInstanceType("m4.large")
 .withMarket("SPOT")
 .withBidPrice("0.25");

InstanceGroupConfig instanceGroupConfigCore = new InstanceGroupConfig()
 .withInstanceCount(4)
 .withInstanceRole("CORE")
 .withInstanceType("m4.large")
 .withMarket("SPOT")
 .withBidPrice("0.03");

InstanceGroupConfig instanceGroupConfigTask = new InstanceGroupConfig()
 .withInstanceCount(2)
```

```
.withInstanceRole("TASK")
.withInstanceType("m4.large")
.withMarket("SPOT")
.withBidPrice("0.10");
```

Flexibilidad de zona de disponibilidad para un clúster de Amazon EMR

Cada una Región de AWS tiene varias ubicaciones aisladas conocidas como zonas de disponibilidad. Cuando lanza una instancia, tiene la opción de especificar una zona de disponibilidad (AZ) en la Región de AWS que está utilizando. La flexibilidad de las zonas de disponibilidad consiste en la distribución de las instancias entre varias AZs. Si una instancia produce un error, puede diseñar la aplicación para que una instancia en otra zona de disponibilidad pueda manejar las solicitudes. Para obtener más información sobre las zonas de disponibilidad, consulta la documentación de regiones y zonas en la Guía del EC2 usuario de Amazon.

La flexibilidad de las instancias consiste en el uso de varios tipos de instancias para satisfacer los requisitos de capacidad. Al expresar la flexibilidad con las instancias, puede utilizar la capacidad agregada en todos los tamaños, familias y generaciones de instancias. Una mayor flexibilidad mejora las posibilidades de encontrar y asignar la cantidad necesaria de capacidad de computación en comparación con un clúster que utiliza un solo tipo de instancia.

La flexibilidad de las instancias y de las zonas de disponibilidad reduce los errores de capacidad insuficiente (ICE) y las interrupciones de spot en comparación con un clúster con un solo tipo de instancia o zona de disponibilidad. Utilice las prácticas recomendadas que se describen aquí para determinar qué instancias debe diversificar una vez que conozca la familia y el tamaño iniciales de las instancias. Este enfoque maximiza la disponibilidad de los grupos de EC2 capacidad de Amazon con una variación mínima de rendimiento y costo.

Ser flexible en cuanto a las zonas de disponibilidad

Se recomienda que configure todas las zonas de disponibilidad para utilizarlas en su nube privada virtual (VPC) y que las seleccione para su clúster de EMR. Los clústeres deben existir en una sola zona de disponibilidad, pero con las flotas de instancias de Amazon EMR, puede seleccionar varias subredes para distintas zonas de disponibilidad. Cuando Amazon EMR lanza el clúster, busca en esas subredes para encontrar las instancias y las opciones de compra que especifique. Cuando aprovisiona un clúster de EMR para varias subredes, su clúster puede acceder a un grupo de EC2 capacidad de Amazon más amplio en comparación con los clústeres de una sola subred.

Si debe priorizar un número determinado de zonas de disponibilidad para usarlas en su nube privada virtual (VPC) para su clúster de EMR, puede aprovechar la capacidad de puntuación de

ubicación puntual con Amazon. EC2 Con la puntuación de ubicación puntual, usted especifica los requisitos de cómputo para sus instancias puntuales y, a continuación, EC2 devuelve las diez zonas de disponibilidad Regiones de AWS o zonas de disponibilidad mejor puntuadas en una escala del 1 al 10. Una puntuación de 10 indica que es muy probable que su solicitud de spot se complete correctamente y una puntuación de 1 indica que es poco probable que esto suceda. Para obtener más información sobre cómo utilizar la puntuación de posicionamiento de Spot, consulta la puntuación de posicionamiento de Spot en la Guía del EC2 usuario de Amazon.

Ser flexible en cuanto a los tipos de instancias

La flexibilidad de las instancias consiste en el uso de varios tipos de instancias para satisfacer los requisitos de capacidad. La flexibilidad de las instancias beneficia al uso de Amazon EC2 Spot y On-Demand Instance. Con las instancias puntuales, la flexibilidad de las instancias permite a Amazon EC2 lanzar instancias desde grupos de capacidad más profundos utilizando datos de capacidad en tiempo real. También predice qué instancias están más disponibles. Esto ofrece menos interrupciones y puede reducir el costo total de una carga de trabajo. Con las instancias bajo demanda, la flexibilidad de las instancias reduce los errores de capacidad insuficiente (ICE) cuando la capacidad total se aprovisiona en un mayor número de grupos de instancias.

En el caso de los clústeres de grupos de instancias, puede especificar hasta 50 tipos de EC2 instancias. En el caso de las flotas de instancias con estrategia de asignación, puedes especificar hasta 30 tipos de EC2 instancias para cada grupo de nodos principal, principal y de tareas. Una gama más amplia de instancias mejora las ventajas de la flexibilidad de las instancias.

Expresar la flexibilidad de las instancias

Tenga en cuenta las siguientes prácticas recomendadas para expresar la flexibilidad de las instancias de su aplicación.

Temas

- Determinar la familia y el tamaño de las instancias
- Incluir instancias adicionales

Determinar la familia y el tamaño de las instancias

Amazon EMR admite varios tipos de instancias para distintos casos de uso. Estos tipos de instancias se enumeran en la documentación Tipos de instancias admitidas con Amazon EMR. Cada tipo de instancia pertenece a una familia de instancias que describe para qué aplicación se ha optimizado el tipo.

Para las cargas de trabajo nuevas, debe compararlas con los tipos de instancias de la familia de uso general, como `m5` o `c5`. Luego, supervise las métricas del sistema operativo y de YARN de Ganglia y Amazon CloudWatch para determinar los cuellos de botella del sistema en momentos de máxima carga. Los cuellos de botella incluyen la CPU, la memoria, el almacenamiento y las operaciones. I/O Después de identificar los cuellos de botella, elija la familia de instancias optimizadas para la computación, optimizadas para memoria, optimizadas para el almacenamiento u otra familia de instancias adecuada para sus tipos de instancias. Para obtener más información, consulte la página Determine la infraestructura adecuada para sus cargas de trabajo de Spark en la guía de prácticas recomendadas de Amazon EMR en. GitHub

A continuación, identifique el contenedor de YARN o el ejecutor de Spark más pequeño que requiere su aplicación. Es el tamaño de instancia más pequeño que cabe en el contenedor y el tamaño mínimo de instancia para el clúster. Utilice esta métrica para determinar las instancias con las que puede diversificar aún más. Una instancia más pequeña permitirá una mayor flexibilidad de instancias.

Para obtener la máxima flexibilidad de instancias, debe utilizar tantas instancias como sea posible. Se recomienda que diversifique con instancias que tengan especificaciones de hardware similares. Esto maximiza el acceso a los grupos de EC2 capacidad con una variación mínima de costos y rendimiento. Diversifique en todos los tamaños. Para ello, priorice primero a AWS Graviton y a las generaciones anteriores. Como regla general, intente ser flexible con al menos 15 tipos de instancias para cada carga de trabajo. Se recomienda que comience con instancias de uso general, optimizadas para la computación u optimizadas para memoria. Estos tipos de instancias ofrecerán la mayor flexibilidad.

Incluir instancias adicionales

Para obtener la máxima diversidad, incluya tipos de instancias adicionales. Priorice primero la flexibilidad de tamaño de la instancia, Graviton y generación. Esto permite el acceso a grupos de EC2 capacidad adicionales con perfiles de costo y rendimiento similares. Si necesita más flexibilidad debido a los ICE o a interrupciones de spot, considere la flexibilidad de variantes y familias. Cada enfoque tiene sus ventajas y desventajas que dependen de su caso de uso y sus requisitos.

- Flexibilidad de tamaño: en primer lugar, diversifique con instancias de diferentes tamaños dentro de la misma familia. Las instancias de la misma familia ofrecen el mismo costo y rendimiento, pero pueden lanzar una cantidad diferente de contenedores en cada host. Por ejemplo, si el tamaño mínimo de ejecutor que necesita es de 2 vCPU y 8 GB de memoria, el tamaño mínimo de instancia es `m5.xlarge`. Para mayor flexibilidad de tamaño, incluya `m5.xlarge`, `m5.2xlarge`, `m5.4xlarge`, `m5.8xlarge`, `m5.12xlarge`, `m5.16xlarge` y `m5.24xlarge`.

- Flexibilidad de Graviton: además del tamaño, puede diversificar con las instancias de Graviton. Las instancias Graviton funcionan con procesadores AWS Graviton2 que ofrecen la mejor relación precio-rendimiento para las cargas de trabajo en la nube de Amazon. EC2 Por ejemplo, con un tamaño de instancia mínimo de `m5.xlarge`, puede incluir `m6g.xlarge`, `m6g.2xlarge`, `m6g.4xlarge`, `m6g.8xlarge` y `m6g.16xlarge` para obtener flexibilidad de Graviton.
- Flexibilidad de generación: al igual que la flexibilidad de Graviton y de tamaño, las instancias de las familias de generaciones anteriores comparten las mismas especificaciones de hardware. Esto se traduce en un perfil de costes y rendimiento similar, con un aumento en el total del EC2 pool de Amazon accesible. Para mayor flexibilidad de generación, incluya `m4.xlarge`, `m4.2xlarge`, `m4.10xlarge` y `m4.16xlarge`.
- Flexibilidad de familias y variantes
  - Capacidad: para optimizar la capacidad, se recomienda la flexibilidad de las instancias en todas las familias de instancias. Las instancias comunes de diferentes familias de instancias tienen grupos de instancias más amplios que pueden ayudar a cumplir los requisitos de capacidad. Sin embargo, las instancias de distintas familias tendrán diferentes proporciones de vCPU y memoria. Esto se traduce en una infrautilización si el tamaño del contenedor de aplicaciones esperado es adecuado para una instancia diferente. Por ejemplo, con `m5.xlarge`, incluya instancias optimizadas para la computación, como `c5`, o instancias optimizadas para memoria, como `r5`, para obtener flexibilidad de familias de instancias.
  - Costo: para optimizar el costo, se recomienda la flexibilidad de las instancias en todas las variantes. Estas instancias tienen la misma proporción de memoria y vCPU que la instancia inicial. La diferencia con la flexibilidad de variantes es que estas instancias tienen grupos de capacidad más pequeños, lo que puede dar como resultado una capacidad adicional limitada o un mayor número de interrupciones de spot. Con `m5.xlarge` como ejemplo, incluya instancias basadas en AMD (`m5a`), instancias basadas en SSD (`m5d`) o instancias optimizadas para la red (`m5n`) para obtener flexibilidad de variantes de instancias.

Configuración de tipos de instancias de clúster de Amazon EMR y prácticas recomendadas para instancias de Spot

Utilice la orientación de esta sección como ayuda para determinar los tipos de instancias, las opciones de compra y la cantidad de almacenamiento para aprovisionar cada tipo de nodo en un clúster de EMR.

¿Qué tipo de instancia debería utilizar?

Hay varias formas de añadir EC2 instancias de Amazon a un clúster. El método que debe elegir depende de si utiliza la configuración de grupos de instancias o la configuración de flotas de instancias para el clúster.

- Grupos de instancias
  - Añada manualmente instancias del mismo tipo a los grupos de instancias de tareas y secundarias existentes.
  - Añada manualmente un grupo de instancias de tareas, que pueden utilizar un tipo de instancia diferente.
  - Configura el escalado automático en Amazon EMR para un grupo de instancias, añadiendo y eliminando instancias automáticamente en función del valor de una CloudWatch métrica de Amazon que especifiques. Para obtener más información, consulte Utilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes.
- Flotas de instancias
  - Añadir una única flota de instancias de tarea.
  - Cambiar la capacidad de destino de las instancias bajo demanda y de spot para las flotas de instancias secundarias y de tareas. Para obtener más información, consulte Planificación y configuración de flotas de instancias para su clúster de Amazon EMR.

Una forma de planificar las instancias del clúster consiste en ejecutar un clúster de prueba con un conjunto representativo de datos de ejemplo y monitorizar la utilización de los nodos del clúster. Para obtener más información, consulte Visualización y monitorización de un clúster de Amazon EMR a medida que realiza su trabajo. Otra forma consiste en calcular la capacidad de las instancias que se estén planeando y comparar dicho valor con el tamaño de los datos.

En general, el tipo de nodo principal, que asigna tareas, no requiere una EC2 instancia con mucha potencia de procesamiento; las EC2 instancias de Amazon para el tipo de nodo principal, que procesan tareas y almacenan datos en HDFS, necesitan potencia de procesamiento y capacidad de almacenamiento; las EC2 instancias de Amazon para el tipo de nodo de tarea, que no almacenan datos, solo necesitan potencia de procesamiento. Para obtener instrucciones sobre las EC2 instancias de Amazon disponibles y su configuración, consulteConfigurar los tipos de EC2 instancias de Amazon para utilizarlos con Amazon EMR.

Las siguientes directrices se aplican a la mayoría de los clústeres de Amazon EMR.

- Hay un límite de vCPU para el número total de EC2 instancias de Amazon bajo demanda que ejecutes en una AWS cuenta por cuenta. Región de AWS Para obtener más información sobre el límite de vCPU y cómo solicitar un aumento del límite para su cuenta, consulte On-Demand Instances en la Guía EC2 del usuario de Amazon para instancias de Linux.
- Generalmente, el nodo principal no tiene grandes requisitos informáticos. Para los clústeres con una gran cantidad de nodos o para los clústeres con aplicaciones que se implementan específicamente en el nodo principal (HueJupyterHub, etc.), es posible que se requiera un nodo principal más grande que ayude a mejorar el rendimiento del clúster. Por ejemplo, considere la posibilidad de utilizar una instancia m5.xlarge para clústeres pequeños (50 nodos o menos) y aumentarla a un tipo de instancia más grande para clústeres más grandes.
- Las necesidades informáticas de los nodos secundarios y de tareas dependen del tipo de procesamiento que realiza la aplicación. Muchos trabajos se pueden ejecutar en tipos de instancia de uso general, que ofrecen un rendimiento equilibrado en términos de CPU, espacio en disco y entrada/salida. Los clústeres con cálculo intensivo podrían beneficiarse de ejecutarse en instancias de uso elevado de CPU, que tienen en proporción más CPU que RAM. Las aplicaciones de base de datos y almacenamiento en memoria caché podrían beneficiarse al ejecutarse en instancias de memoria elevada. Las aplicaciones con uso intensivo de red y uso intensivo de CPU como análisis, NLP y machine learning pueden beneficiarse de las instancias de computación en clúster, que ofrecen proporcionalmente recursos de CPU elevada y mayor rendimiento de red.
- Si distintas fases del clúster tienen diferentes necesidades de capacidad, puede empezar con un pequeño número de nodos secundarios y aumentar o reducir el número de nodos de tareas para satisfacer los requisitos de capacidad variable del flujo de trabajo.
- La cantidad de datos que puede procesar depende de la capacidad de los nodos secundarios y del tamaño de los datos como entrada, durante el procesamiento y como salida. Los conjuntos de datos entrantes, intermedios y salientes residen en el clúster durante el procesamiento.

¿Cuándo se deben utilizar las instancias de spot?

Al lanzar un clúster en Amazon EMR, puede elegir lanzar instancias principales, básicas y de tarea en instancias de spot. Dado que cada tipo de grupo de instancias desempeña un papel diferente en el clúster, hay distintas consecuencias al lanzar cada tipo de nodo en instancias de spot. No se puede cambiar una opción de compra de instancias mientras se ejecuta un clúster. Para cambiar de instancias bajo demanda a instancias de spot o viceversa para el nodo principal y los nodos básicos, debe terminar el clúster y lanzar uno nuevo. Para los nodos de tareas, puede lanzar un nuevo grupo de instancias de tareas o una nueva flota de instancias y eliminar la anterior.

Temas

- Configuración de Amazon EMR para evitar errores en los trabajos debido a la terminación de instancias de spot de los nodos de tarea
- Nodo principal en una instancia de spot
- Nodos básicos en instancias de spot
- Nodos de tarea en instancias de spot
- Configuraciones de instancias para escenarios de aplicaciones

Configuración de Amazon EMR para evitar errores en los trabajos debido a la terminación de instancias de spot de los nodos de tarea

Dado que las instancias de spot se utilizan a menudo para ejecutar nodos de tarea, Amazon EMR tiene una funcionalidad predeterminada para programar trabajos de YARN, de modo que los trabajos en ejecución no presenten errores cuando los nodos de tarea que se ejecutan en las instancias de spot se terminen. Para ello, Amazon EMR permite que los procesos maestros de la aplicación se ejecuten únicamente en los nodos principales. El proceso maestro de la aplicación controla los trabajos en ejecución y debe mantenerse activo durante toda la vida del trabajo.

La versión 5.19.0 y posteriores de Amazon EMR utilizan la característica integrada de etiquetas de nodo YARN para lograrlo. (Las versiones anteriores utilizaban una revisión de código). Las propiedades en las clasificaciones de configuración `yarn-site` y `capacity-scheduler` se ajustan de forma predeterminada para que capacity-scheduler y fair-scheduler de YARN utilicen las etiquetas de nodo. Amazon EMR etiqueta automáticamente los nodos principales con la etiqueta `CORE` y establece las propiedades para que los maestros de la aplicación se programen únicamente en los nodos con la etiqueta CORE. La modificación manual de las propiedades relacionadas en las clasificaciones de configuración yarn-site y capacity-scheduler o directamente en los archivos XML asociados podría interrumpir esta característica o modificar esta funcionalidad.

Amazon EMR configura las siguientes propiedades y valores de forma predeterminada. Actúe con precaución al configurar estas propiedades.

> **Note**
>
> A partir de la serie de versiones 6.x de Amazon EMR, la característica de etiquetas de nodo YARN está desactivada de forma predeterminada. De forma predeterminada, los procesos principales de la aplicación se pueden ejecutar tanto en nodos básicos como en nodos de

tarea. Puede habilitar la función de etiquetas de nodo YARN configurando las siguientes propiedades:

- `yarn.node-labels.enabled: true`
- `yarn.node-labels.am.default-node-label-expression: 'CORE'`

- yarn-site (yarn-site.xml) en todos los nodos
  - `yarn.node-labels.enabled: true`
  - `yarn.node-labels.am.default-node-label-expression: 'CORE'`
  - `yarn.node-labels.fs-store.root-dir: '/apps/yarn/nodelabels'`
  - `yarn.node-labels.configuration-type: 'distributed'`
- yarn-site (yarn-site.xml) en los nodos principal y básicos
  - `yarn.nodemanager.node-labels.provider: 'config'`
  - `yarn.nodemanager.node-labels.provider.configured-node-partition: 'CORE'`
- capacity-scheduler (capacity-scheduler.xml) en todos los nodos
  - `yarn.scheduler.capacity.root.accessible-node-labels: '*'`
  - `yarn.scheduler.capacity.root.accessible-node-labels.CORE.capacity: 100`
  - `yarn.scheduler.capacity.root.default.accessible-node-labels: '*'`
  - `yarn.scheduler.capacity.root.default.accessible-node-labels.CORE.capacity: 100`

Nodo principal en una instancia de spot

El nodo principal controla y dirige el clúster. Cuando se termina, el clúster finaliza, por lo que solo debe lanzar el nodo principal como una instancia de spot si está ejecutando un clúster que se acepta que termine de forma repentina. Podría ser el caso si está probando una aplicación nueva, tiene un clúster que guarda periódicamente datos en un almacén externo como Amazon S3 o está ejecutando un clúster donde el costo es más importante que garantizar que este se complete.

Cuando se lanza el grupo de instancias principales como instancia de spot, el clúster no se inicia hasta que se completa la solicitud de instancia de spot. Esto es algo que debe considerar a la hora de seleccionar el precio de spot máximo.

Solo puede agregar un nodo principal de instancias de spot al lanzar el clúster. No se puede agregar ni eliminar nodos principales de un clúster en ejecución.

Normalmente, solo se ejecutaría el nodo principal como instancia de spot si se ejecuta todo el clúster (todos los grupos de instancia) como instancias de spot.

Nodos básicos en instancias de spot

Los nodos secundarios procesan los datos y almacenan información mediante HDFS. La terminación de una instancia secundaria conlleva el riesgo de pérdida de datos. Por este motivo, solo debe ejecutar los nodos secundarios en instancias de spot cuando sea admisible la pérdida parcial de datos de HDFS.

Cuando se lanza el grupo de instancias básicas como instancias de spot, Amazon EMR espera hasta que se puedan aprovisionar todas las instancias básicas solicitadas antes de lanzar el grupo de instancias. En otras palabras, si solicitas seis EC2 instancias de Amazon y solo hay cinco disponibles al precio de subasta máximo o inferior, el grupo de instancias no se lanzará. Amazon EMR seguirá esperando hasta que las seis EC2 instancias de Amazon estén disponibles o hasta que cancele el clúster. Puede cambiar el número de instancias de spot de un grupo de instancias secundario para añadir capacidad a un clúster en ejecución. Para obtener más información sobre cómo trabajar con los grupos de instancias y cómo funcionan las instancias de spot con las flotas de instancias, consulte the section called “Configurar flotas de instancias o grupos de instancias”.

Nodos de tarea en instancias de spot

Los nodos de tareas procesan datos pero no guardan datos persistentes en HDFS. Si se terminan porque el precio de spot ha superado su precio de spot máximo, no se pierden los datos y el efecto sobre el clúster es mínimo.

Al lanzar uno o varios grupos de instancias de tarea como instancias de spot, Amazon EMR aprovisiona tantos nodos de tarea como puede con su precio de spot máximo. Esto significa que, si solicita un grupo de instancias de tarea con seis nodos y solo hay cinco instancias de spot disponibles a su precio de spot máximo o por debajo de este, Amazon EMR lanza el grupo de instancias con cinco nodos y agrega el sexto más tarde, si es posible.

El lanzamiento de grupos de instancias de tareas como instancias de spot es una forma estratégica de ampliar la capacidad del clúster minimizando los costos. Si lanza los grupos de instancias principales y básicas como instancias bajo demanda, su capacidad está garantizada para la ejecución del clúster. Puede añadir instancias de tarea a los grupos de instancias de tarea según sea necesario para gestionar los picos de tráfico o acelerar el procesamiento de datos.

Puede añadir o eliminar nodos de tareas mediante la consola o AWS CLI la API. También puede añadir grupos de tareas adicionales, pero no puede quitar un grupo de tareas después de haberlo creado.

Configuraciones de instancias para escenarios de aplicaciones

La siguiente tabla es una referencia rápida para las opciones de compra de tipos de nodos y las configuraciones que suelen ser adecuadas para los distintos escenarios de aplicaciones. Haga clic en el enlace correspondiente para ver más información sobre cada escenario.

Escenario de aplicaciones	Opción de compra del nodo principal	Opción de compra de los nodos básicos	Opción de compra de nodos de tarea
Clústeres en ejecución prolongada y almacenamientos de datos	Bajo demanda	Bajo demanda o combinación de flotas de instancias	Spot o combinación de flotas de instancias
Cargas de trabajo dirigidas por costos	Spot	Spot	Spot
Cargas de trabajo críticas para datos	Bajo demanda	Bajo demanda	Spot o combinación de flotas de instancias
Prueba de aplicación	Spot	Spot	Spot

Existen varios escenarios en las que las instancias de spot son útiles para ejecutar un clúster de Amazon EMR.

Clústeres en ejecución prolongada y almacenamientos de datos

Si ejecuta un clúster de Amazon EMR persistente que tiene una variación predecible de capacidad informática, como un almacenamiento de datos, puede gestionar los picos de demanda a menor costo con las instancias de spot. Puede lanzar los grupos de instancias principales y básicas como instancias bajo demanda para gestionar la capacidad normal y lanzar los grupos de instancias de tarea como instancias de spot para gestionar los requisitos de picos de carga.

Cargas de trabajo dirigidas por costos

Si está ejecutando clústeres transitorios para los que un costo inferior es más importante que el tiempo de finalización y es aceptable la pérdida parcial de trabajo, puede ejecutar todo el clúster (grupos de instancias principales, básicas y de tarea) como instancias de spot para beneficiarse de los mayores ahorros de costos.

Cargas de trabajo críticas para datos

Si está ejecutando un clúster para el que un costo inferior es más importante que el tiempo de finalización, pero no es aceptable la pérdida parcial de trabajo, lance los grupos de instancias principales y básicas como instancias bajo demanda y compleméntelas con uno o varios grupos de instancias de tarea de instancias de spot. La ejecución de los grupos de instancias principales y básicas como instancias bajo demanda garantiza que los datos se almacenen en HDFS y que el clúster se proteja de la terminación debido a fluctuaciones en el mercado de spot, además de proporcionar un ahorro de costos que supone la ejecución de los grupos de instancias de tarea como instancias de spot.

Prueba de aplicación

Cuando se prueba una nueva aplicación para prepararla para su lanzamiento en un entorno de producción, se puede ejecutar todo el clúster (grupos de instancias principales, básicas y de tarea) como instancias de spot para reducir los costos de las pruebas.

Cálculo de la capacidad de HDFS requerida de un clúster

La cantidad de almacenamiento HDFS disponible para su clúster depende de los siguientes factores:

- El número de EC2 instancias de Amazon utilizadas para los nodos principales.
- La capacidad del almacén de EC2 instancias de Amazon para el tipo de instancia utilizado. Para obtener más información sobre los volúmenes de los almacenes de instancias, consulta Amazon EC2 Instance Store en la Guía del EC2 usuario de Amazon.
- El número y el tamaño de los volúmenes de Amazon EBS asociados a los nodos básicos.
- Un factor de replicación, que tiene en cuenta la cantidad de bloques de datos que se almacenan en HDFS para redundancia similar a RAID. De forma predeterminada, el factor de replicación es tres para un clúster de 10 o más nodos secundarios, dos para un clúster de 4-9 nodos secundarios y uno para un clúster de 3 nodos o menos.

Para calcular la capacidad de HDFS de un clúster, por cada nodo principal, agregue la capacidad del volumen del almacén de instancias a la capacidad de almacenamiento de Amazon EBS (si se utiliza). Multiplique el resultado por el número de nodos secundarios y, a continuación, divida el total por el factor de replicación que depende del número de nodos secundarios. Por ejemplo, un clúster con 10 nodos básicos de tipo i2.xlarge, que tiene 800 GB de almacenamiento de instancias y no tiene ningún volumen de Amazon EBS asociado, tiene un total de aproximadamente 2666 GB disponibles para HDFS (10 nodos x 800 GB ÷ 3 de factor de replicación).

Si el valor de capacidad de HDFS calculado es inferior a sus datos, puede aumentar la cantidad de almacenamiento de HDFS de las siguientes formas:

- Creación de un clúster con volúmenes de Amazon EBS adicionales o adición de grupos de instancias con volúmenes de Amazon EBS asociados a un clúster existente
- Agregar más nodos secundarios
- Elegir un tipo de EC2 instancia de Amazon con mayor capacidad de almacenamiento
- Uso de la compresión de datos
- Cambio de la configuración de Hadoop para reducir el factor de replicación

La reducción del factor de replicación se debería utilizar con precaución, ya que reduce la redundancia de los datos de HDFS y la capacidad del clúster de recuperarse frente a bloques de HDFS perdidos o dañados.

Configuración del registro y la depuración de un clúster de Amazon EMR

Una de las cosas que decidir a la hora de planificar el clúster es la cantidad de soporte de depuración que tiene que hacer que esté disponible. Al desarrollar por primera vez la aplicación de procesamiento de datos, le recomendamos probar la aplicación en un clúster que procese un subconjunto pequeño, pero representativo, de sus datos. Al hacerlo, probablemente querrá aprovechar todas las herramientas de depuración que ofrece Amazon EMR, como el archivado de archivos de registro en Amazon S3.

Cuando haya terminado el desarrollo y tenga la aplicación de procesamiento de datos a pleno rendimiento, puede optar por reducir la depuración. Al hacerlo puede ahorrar el costo de almacenamiento del conjunto de los archivos de registro en Amazon S3 y reducir la carga de procesamiento en el clúster, ya que no es necesario escribir el estado en Amazon S3. El

inconveniente, por supuesto, es que si algo va mal, tendrá menos herramientas disponibles para investigar el problema.

Archivos de registro predeterminados

De forma predeterminada, cada clúster escribe archivos de registro en el nodo principal. Se escriben en el directorio `/mnt/var/log/`. Puede acceder a ellos a través de SSH para conectarse al nodo principal tal como se describe en Conexión al nodo principal del clúster de Amazon EMR mediante SSH. Amazon EMR recopila determinados registros de sistemas y aplicaciones generados por los daemons de Amazon EMR y otros procesos de Amazon EMR para garantizar la eficacia de las operaciones de servicio.

> **Note**
>
> Si utiliza la versión 6.8.0 o anterior de Amazon EMR, los archivos de registro se guardan en Amazon S3 durante la terminación del clúster, por lo que no podrá acceder a los archivos de registro una vez que termine el nodo principal. Las versiones 6.9.0 y posteriores de Amazon EMR archivan los registros en Amazon S3 durante la reducción vertical del clúster, de modo que los archivos de registro generados en él persisten incluso después de terminar el nodo.

No tiene que habilitar nada para que los archivos de registro se escriban en el nodo principal. Este es el comportamiento predeterminado de Amazon EMR y Hadoop.

Un clúster genera varios tipos de archivos de registro, incluidos:

- Registros de paso: estos registros se generan mediante el servicio de Amazon EMR y contienen información acerca del clúster y los resultados de cada paso. Los archivos de registro se almacenan en el directorio `/mnt/var/log/hadoop/steps/` en el nodo principal. Cada paso registra sus resultados en un subdirectorio numerado independiente: `/mnt/var/log/hadoop/steps/s-`*`stepId1`*`/` para el primer paso, `/mnt/var/log/hadoop/steps/s-`*`stepId2`*`/` para el segundo paso, y así sucesivamente. Los identificadores de paso de 13 caracteres (p. ej. stepId1, stepId2) son exclusivos de un clúster.
- Registros de componentes de Hadoop e YARN: los registros de los componentes asociados a Apache YARN y MapReduce, por ejemplo, se encuentran en carpetas independientes. `/mnt/var/log` Las ubicaciones de archivo de registro para los componentes de Hadoop en `/mnt/var/log` son las siguientes: hadoop-hdfs, hadoop-mapreduce, hadoop-httpfs y hadoop-yarn. El

hadoop-state-pusher directorio es para el resultado del proceso de introducción de estados de Hadoop.

- Registros de acción de arranque: si su trabajo utiliza acciones de arranque, se registran los resultados de estas acciones. Los archivos de registro se almacenan en/mnt/var/log/bootstrap-actions/ en el nodo principal. Cada acción de arranque registra sus resultados en un subdirectorio numerado independiente: `/mnt/var/log/bootstrap-actions/1/` para la primera acción de arranque, `/mnt/var/log/bootstrap-actions/2/` para la segunda acción de arranque y así sucesivamente.
- Registros de estado de instancia: estos registros proporcionan información sobre la CPU, el estado de la memoria y los subprocesos del recolector de elementos no utilizados del nodo. Los archivos de registro se almacenan en `/mnt/var/log/instance-state/` en el nodo principal.

Archivar archivos de registro en Amazon S3

> **Note**
>
> Actualmente, no puede utilizar la agregación de registros para Amazon S3 con la utilidad `yarn logs`.

Las versiones 6.9.0 y posteriores de Amazon EMR archivan los registros en Amazon S3 durante la reducción vertical del clúster, de modo que los archivos de registro generados en él persisten incluso después de terminar el nodo. Este comportamiento se habilita de forma automática, por lo que no es necesario que haga nada para activarlo. Para las versiones 6.8.0 y anteriores de Amazon EMR, puede configurar un clúster para archivar periódicamente los archivos de registro almacenados en el nodo principal en Amazon S3. Esto garantiza que los archivos de registro estén disponibles después de que el clúster se termine, ya sea a través de un apagado normal o debido a un error. Amazon EMR archiva los archivos de registro en Amazon S3 en intervalos de 5 minutos.

Para que los archivos de registro se archiven en Amazon S3 en el caso de las versiones 6.8.0 y anteriores de Amazon EMR, debe habilitar esta característica al lanzar el clúster. Puede hacerlo utilizando la consola, la CLI o la API. De manera predeterminada, los clústeres lanzados con la consola tienen habilitado el archivado de registros. Para clústeres lanzados con la CLI o la API, el registro en Amazon S3 se debe habilitar manualmente.

Console

Para archivar los archivos de registro en Amazon S3 con la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Registros de clúster, seleccione la casilla Publicar los registros específicos del clúster en Amazon S3.
4. En el campo Ubicación de Amazon S3, escriba (o busque) una ruta de Amazon S3 para almacenar sus registros. Si escribe el nombre de una carpeta que no existe en el bucket, Amazon S3 la creará.

   Al establecer este valor, Amazon EMR copia los archivos de registro de las EC2 instancias del clúster a Amazon S3. Esto evita que los archivos de registro se pierdan cuando el clúster finaliza y terminan las instancias que EC2 alojan el clúster. Estos registros resultan útiles en la solución de problemas. Para más información, consulte Ver archivos de registro.
5. De manera opcional, seleccione la casilla Cifrar los registros específicos del clúster. A continuación, seleccione una AWS KMS clave de la lista, introduzca un ARN de clave o cree una clave nueva. Esta opción solo está disponible con la versión 5.30.0 y versiones posteriores de Amazon EMR, excluyendo la versión 6.0.0. Para usar esta opción, añada permisos a su perfil AWS KMS de EC2 instancia y rol de Amazon EMR. Para obtener más información, consulte Para cifrar los archivos de registro almacenados en Amazon S3 con una clave administrada por el cliente de AWS KMS.
6. Elija cualquier otra opción que se aplique a su clúster.
7. Para lanzar el clúster, elija Crear clúster.

CLI

Para archivar archivos de registro en Amazon S3 con el AWS CLI

Para archivar archivos de registro en Amazon S3 mediante el AWS CLI, escriba el `create-cluster` comando y especifique la ruta de registro de Amazon S3 mediante el `--log-uri` parámetro.

1. Para registrar archivos en Amazon S3, escriba el siguiente comando y *myKey* sustitúyalo por el nombre del par de EC2 claves.

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.10.0 --log-
uri s3://DOC-EXAMPLE-BUCKET/logs --applications Name=Hadoop Name=Hive Name=Pig
 --use-default-roles --ec2-attributes KeyName=myKey --instance-type m5.xlarge --
instance-count 3
```

2. Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos básicos. Todos los nodos utilizarán el tipo de instancia especificado en el comando.

> **Note**
>
> Si no ha creado anteriormente el rol de servicio Amazon EMR y el perfil de EC2 instancia predeterminados, introduzca `aws emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Para cifrar los archivos de registro almacenados en Amazon S3 con una clave administrada por el cliente de AWS KMS

Con Amazon EMR versión 5.30.0 y versiones posteriores (excepto Amazon EMR 6.0.0), puede cifrar los archivos de registro almacenados en Amazon S3 con una clave de KMS gestionada por el cliente. AWS Para habilitar esta opción en la consola, siga los pasos descritos en Archivar archivos de registro en Amazon S3. Su perfil de EC2 instancia de Amazon y su función de Amazon EMR deben cumplir los siguientes requisitos previos:

- El perfil de EC2 instancia de Amazon utilizado para tu clúster debe tener permiso de usokms`:GenerateDataKey`.
- El rol de Amazon EMR utilizado para el clúster debe tener permiso para utilizar `kms:DescribeKey`.
- El perfil de la EC2 instancia de Amazon y la función de Amazon EMR se deben añadir a la lista de usuarios clave de la clave gestionada por el cliente de AWS KMS especificada, como se demuestra en los siguientes pasos:
  1. Abre la consola AWS Key Management Service (AWS KMS) en https://console.aws.amazon.com/kms.

2. Para cambiar la AWS región, utilice el selector de regiones situado en la esquina superior derecha de la página.
3. Seleccione el alias de la clave de KMS que desee modificar.
4. En la página de detalles de la clave, en Key Users (Usuarios de claves), seleccione Add (Añadir).
5. En el cuadro de diálogo Añadir usuarios clave, seleccione su perfil de EC2 instancia de Amazon y su función de Amazon EMR.
6. Elija Agregar.

- También debe configurar la clave KMS para permitir que los directores `elasticmapreduce.amazonaws.com` de servicio puedan`kms:GenerateDataKey`, `persistentappui.elasticmapreduce.amazonaws.com` `kms:GenerateDataKeyWithoutPlaintext` y. `kms:Decrypt` Esto permite a EMR leer y escribir registros cifrados con la clave KMS en el almacenamiento gestionado. El rol de IAM del usuario debe tener permiso para usar `kms:GenerateDataKey` y. `kms:Decrypt`

```
{
   "Sid": "Allow User Role to use KMS key",
   "Effect": "Allow",
   "Principal": {
        "AWS": "User Role"
    },
    "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey"
   ],
    "Resource": "*",
    "Condition": {
        "StringLike": {
            "kms:EncryptionContext:aws:elasticmapreduce:clusterId": "j-*",
           "kms:ViaService": "elasticmapreduce.region.amazonaws.com"
       }
    }
},
{
    "Sid": "Allow Persistent APP UI to validate KMS key for write",
    "Effect": "Allow",
    "Principal":{
        "Service": [
            "elasticmapreduce.amazonaws.com"
```

```
        ]
     },
     "Action": [
       "kms:GenerateDataKeyWithoutPlaintext"
      ],
     "Resource": "*",
     "Condition": {
        "StringLike": {
            "aws:SourceArn": "arn:aws:elasticmapreduce:region:account:cluster/j-*",
            "kms:EncryptionContext:aws:elasticmapreduce:clusterId": "j-*"
        }
     }
},
{
    "Sid": "Allow Persistent APP UI to Write/Read Logs",
    "Effect": "Allow",
    "Principal":{
        "Service": [
            "persistentappui.elasticmapreduce.amazonaws.com",
            "elasticmapreduce.amazonaws.com"
        ]
     },
     "Action": [
       "kms:Decrypt",
       "kms:GenerateDataKey"
     ],
     "Resource": "*",
     "Condition": {
        "StringLike": {
            "aws:SourceArn": "arn:aws:elasticmapreduce:region:account:cluster/j-*",
            "kms:EncryptionContext:aws:elasticmapreduce:clusterId": "j-*",
            "kms:ViaService": "s3.region.amazonaws.com"
        }
     }
}
```

Como práctica recomendada de seguridad, le recomendamos que añada las `aws:SourceArn` condiciones `kms:EncryptionContext` y. Estas condiciones ayudan a garantizar que Amazon EMR solo utilice la clave EC2 y solo para los registros generados a partir de trabajos que se ejecutan en un clúster específico.

Para obtener más información, consulte las funciones de servicio de IAM utilizadas por Amazon EMR y el uso de políticas clave en AWS la guía para desarrolladores de Key Management Service.

Para agregar registros en Amazon S3 con la AWS CLI

> **Note**
>
> No puede utilizar actualmente la agregación de registros con la utilidad `yarn logs`. Solo puede utilizar la agregación compatible con este procedimiento.

La agregación de registros (Hadoop 2.x) compila los registros desde todos los contenedores para una aplicación individual en un único archivo. Para habilitar la agregación de registros en Amazon S3 mediante el AWS CLI, utilice una acción de arranque al lanzar el clúster para habilitar la agregación de registros y especificar el depósito en el que se almacenarán los registros.

- Para habilitar la agregación de registros, cree el siguiente archivo de configuración llamado `myConfig.json`, que contiene lo siguiente:

```
[
  {
    "Classification": "yarn-site",
    "Properties": {
      "yarn.log-aggregation-enable": "true",
      "yarn.log-aggregation.retain-seconds": "-1",
      "yarn.nodemanager.remote-app-log-dir": "s3:\/\/DOC-EXAMPLE-BUCKET\/logs"
    }
  }
]
```

  Escriba el siguiente comando y *myKey* sustitúyalo por el nombre de su EC2 key pair. Además, puede sustituir cualquiera de los textos en rojo por sus propias configuraciones.

```
aws emr create-cluster --name "Test cluster" \
--release-label emr-7.10.0 \
--applications Name=Hadoop \
--use-default-roles \
--ec2-attributes KeyName=myKey \
--instance-type m5.xlarge \
--instance-count 3 \
```

```
--configurations file://./myConfig.json
```

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos básicos. Todos los nodos utilizarán el tipo de instancia especificado en el comando.

> **Note**
>
> Si no ha creado anteriormente el rol de servicio EMR y el perfil de EC2 instancia predeterminados, ejecute `aws emr create-default-roles` para crearlos antes de ejecutar el `create-cluster` subcomando.

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte la Referencia de AWS CLI comandos.

Localización de registros

La siguiente lista incluye todos los tipos de registro y sus ubicaciones en Amazon S3. Puede utilizarlos para solucionar problemas de Amazon EMR.

Registros de pasos

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/steps/<step-id>/
```

Registros de aplicaciones

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/containers/
```

Esta ubicación incluye el contenedor `stderr` y los registros `stdout`, `directory.info`, `prelaunch.out` y `launch_container.sh`.

Registros del administrador de recursos

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<leader-instance-id>/
applications/hadoop-yarn/
```

HDFS de Hadoop

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<all-instance-id>/
applications/hadoop-hdfs/
```

Esta ubicación incluye NameNode DataNode, y los TimelineServer registros de YARN.

Registros del administrador de nodos

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<all-instance-id>/
applications/hadoop-yarn/
```

Registros de estado de la instancia

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<all-instance-id>/daemons/
instance-state/
```

Registros de aprovisionamiento de Amazon EMR

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<leader-instance-id>/
provision-node/*
```

Registros de Hive

```
s3://DOC-EXAMPLE-LOG-BUCKET/<cluster-id>/node/<leader-instance-id>/
applications/hive/*
```

- Para encontrar los registros de Hive en su clúster, elimine el asterisco (`*`) y agregue `/var/log/hive/` al enlace anterior.
- Para encontrar HiveServer 2 registros, elimine el asterisco (`*`) y añádalo `var/log/hive/hiveserver2.log` al enlace anterior.
- Para encontrar los registros de HiveCLI, elimine el asterisco (`*`) y agregue `/var/log/hive/user/hadoop/hive.log` al enlace anterior.
- Para encontrar los registros de Hive Metastore Server, elimine el asterisco (`*`) y agregue `/var/log/hive/user/hive/hive.log` al enlace anterior.

Si el error se encuentra en el nodo principal o de tarea de su aplicación de Tez, proporcione los registros del contenedor de Hadoop correspondiente.

Etiquetado y clasificación de los recursos del clúster de Amazon EMR

Puede ser conveniente clasificar los AWS recursos de diferentes maneras; por ejemplo, por propósito, propietario o entorno. Puede conseguirlo en Amazon EMR asignando metadatos personalizados a los clústeres de Amazon EMR mediante etiquetas. Una etiqueta consta de una clave y un valor, ambos definidos por el usuario. En el caso de Amazon EMR, el clúster es el nivel de

recursos que puede etiquetar. Por ejemplo, podría definir un conjunto de etiquetas para los clústeres de su cuenta que le ayude a realizar un seguimiento de cada propietario del clúster o identificar un clúster de producción frente a un clúster de pruebas. Le recomendamos que cree un conjunto de etiquetas coherente para satisfacer los requisitos de su organización.

Al añadir una etiqueta a un clúster de Amazon EMR, la etiqueta también se propaga a cada EC2 instancia de Amazon activa asociada al clúster. Del mismo modo, cuando eliminas una etiqueta de un clúster de Amazon EMR, esa etiqueta se elimina de cada instancia de Amazon EC2 activa asociada.

> **Important**
>
> Utilice la consola o CLI de Amazon EMR para gestionar las etiquetas en las EC2 instancias de Amazon que forman parte de un clúster en lugar de la EC2 consola o la CLI de Amazon, ya que los cambios que realice en Amazon EC2 no se sincronizan de nuevo con el sistema de etiquetado de Amazon EMR.

Para identificar una EC2 instancia de Amazon que forma parte de un clúster de Amazon EMR, busque las siguientes etiquetas de sistema. En este ejemplo, *CORE* es el valor del rol del grupo de instancias y *j-12345678* es un ejemplo del valor del identificador del flujo de trabajo (clúster):

- aws:elasticmapreduce: = instance-group-role *CORE*
- aws:elasticmapreduce: = job-flow-id *j-12345678*

> **Note**
>
> Amazon EMR y Amazon EC2 interpretan las etiquetas como una cadena de caracteres sin significado semántico.

Puede trabajar con etiquetas mediante AWS Management Console la CLI y la API.

Puede agregar etiquetas al crear un clúster nuevo de Amazon EMR y puede agregar, editar o eliminar etiquetas de un clúster de Amazon EMR en ejecución. Editar una etiqueta es un concepto que se aplica a la consola de Amazon EMR; sin embargo, con la CLI y la API, para editar una etiqueta, se elimina la etiqueta antigua y se agrega una nueva. Puede editar las claves y los valores de las etiquetas y también puede eliminar etiquetas de un recurso en cualquier momento

que se ejecuta un clúster. Sin embargo, no puede añadir, editar o suprimir etiquetas desde un clúster terminado o instancias terminadas que se asociaron anteriormente con un clúster que sigue activo. Además, puede establecer el valor de una etiqueta como una cadena vacía, pero no puede establecer el valor de una etiqueta como nulo.

Si utilizas AWS Identity and Access Management (IAM) con tus EC2 instancias de Amazon para permisos basados en recursos por etiqueta, tus políticas de IAM se aplican a las etiquetas que Amazon EMR propaga a las instancias de Amazon de un clúster. EC2 Para que las etiquetas EMR de Amazon se propaguen a tus EC2 instancias de Amazon, tu política de IAM para Amazon EC2 debe permitir permisos para llamar a Amazon y. EC2 CreateTags DeleteTags APIs Además, las etiquetas propagadas pueden afectar a los permisos basados en recursos EC2 de Amazon. Las etiquetas propagadas a Amazon se EC2 pueden leer como condiciones en tu política de IAM, al igual que otras etiquetas de Amazon EC2 . Tenga en cuenta la política de IAM a la hora de agregar etiquetas a las clústeres de Amazon EMR para evitar que los usuarios tengan permisos incorrectos para un clúster. Para evitar problemas, asegúrese de que sus políticas de IAM no incluyan condiciones de etiquetas que también planee utilizar en sus clústeres de Amazon EMR. Para obtener más información, consulte Controlar el acceso a EC2 los recursos de Amazon.

Restricciones que se aplican al etiquetado de recursos en Amazon EMR

Se aplican las siguientes restricciones básicas a las etiquetas:

- Las restricciones que se aplican a EC2 los recursos de Amazon también se aplican a Amazon EMR. Para obtener más información, consulte https://docs.aws.amazon.com/AWSEC2/latest/UserGuide/Using_Tags.html#tag-restrictions.
- No utilice el `aws:` prefijo en los nombres y valores de las etiquetas porque está reservado para su AWS uso. Además, no puede editar o eliminar los nombres y valores de etiquetas que tienen este prefijo.
- No puede cambiar o editar etiquetas en un clúster terminado.
- Un valor de etiqueta puede ser una cadena vacía, pero no nulo. Además, una clave de etiqueta no puede ser una cadena vacía.
- Las claves y los valores pueden contener caracteres alfabéticos en cualquier idioma, cualquier carácter numérico, espacios en blanco, separadores invisibles y los siguientes símbolos: _ . : / = + - @

Para obtener más información sobre el etiquetado mediante el AWS Management Console, consulte Trabajar con etiquetas en la consola en la Guía del EC2 usuario de Amazon. Para obtener más

información sobre el etiquetado mediante la EC2 API de Amazon o la línea de comandos, consulte la descripción general de la API y la CLI en la Guía del EC2 usuario de Amazon.

Etiquetado de recursos de Amazon EMR para facturación

Puede usar etiquetas para organizar su AWS factura y reflejar su propia estructura de costos. Para ello, regístrese para recibir la factura de su AWS cuenta con los valores clave de las etiquetas incluidos. A continuación, puede organizar su información de facturación por valores clave de etiqueta, para ver el costo de los recursos combinados. Aunque Amazon EMR y Amazon EC2 tienen estados de facturación diferentes, las etiquetas de cada clúster también se colocan en cada instancia asociada, de modo que puede utilizarlas para vincular los costes relacionados con Amazon EMR y Amazon. EC2

Por ejemplo, puede etiquetar varios recursos con un nombre de aplicación específico y luego organizar su información de facturación para ver los costos totales de la aplicación en distintos servicios. Para más información, consulte Etiquetado y asignación de costos en la Guía del usuario de AWS Billing .

Cómo añadir etiquetas a un clúster de Amazon EMR

También puede agregar etiquetas a un clúster al momento de crearlo.

Console

Para agregar etiquetas al crear un clúster con la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Etiquetas, elija Agregar etiqueta nueva. Especifique una etiqueta en el campo Clave. Si lo desea, especifique una etiqueta en el campo Valor.
4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para añadir etiquetas al crear un clúster con el AWS CLI

El siguiente ejemplo muestra cómo añadir una etiqueta a un nuevo clúster con la AWS CLI. Para añadir etiquetas al crear un clúster, escriba el subcomando `create-cluster` con el parámetro `--tags`.

- Para añadir una etiqueta *`costCenter`* con un valor clave *`marketing`* al crear un clúster, escriba el siguiente comando y *`myKey`* sustitúyalo por el nombre del par de EC2 claves.

```
aws emr create-cluster --name "Test cluster" --release-label emr-4.0.0 --
applications Name=Hadoop Name=Hive Name=Pig --tags "costCenter=marketing" --
use-default-roles --ec2-attributes KeyName=myKey --instance-type m5.xlarge --
instance-count 3
```

Cuando especifica el recuento de instancias sin utilizar el parámetro `--instance-groups`, se lanza un nodo principal único y el resto de las instancias se lanzan como nodos secundarios. Todos los nodos utilizarán el tipo de instancia especificado en el comando.

> **Note**
>
> Si no ha creado anteriormente el rol de servicio EMR y el perfil de EC2 instancia predeterminados, escriba aws `emr create-default-roles` para crearlos antes de escribir el `create-cluster` subcomando.

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

También puede añadir etiquetas a un clúster existente.

Console

Para agregar etiquetas a un clúster existente utilizando la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.

2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. En la pestaña Etiquetas de la página de detalles del clúster, seleccione Administrar etiquetas. Especifique una etiqueta en el campo Clave. Si lo desea, especifique una etiqueta en el campo Valor.
4. Seleccione Guardar cambios. La pestaña Etiquetas se actualiza con el nuevo número de etiquetas que tiene en el clúster. Por ejemplo, si ahora tiene dos etiquetas, la etiqueta de su pestaña será Etiquetas (2).

AWS CLI

Para añadir etiquetas a un clúster en ejecución con la AWS CLI

- Escriba el subcomando `add-tags` con el parámetro `--tag` para asignar etiquetas a un ID de clúster. Puede encontrar el ID del clúster mediante la consola o el comando `list-clusters`. El subcomando `add-tags` actualmente solo acepta un ID de recurso.

  Por ejemplo, para añadir dos etiquetas a un clúster en ejecución, una con una clave denominada *`costCenter`* con un valor de *`marketing`* y otra *`other`* con un valor de*`accounting`*, introduzca el siguiente comando y sustitúyalo por *`j-KT4XXXXXXXX1NM`* el ID del clúster.

  ```
  aws emr add-tags --resource-id j-KT4XXXXXXXX1NM --tag "costCenter=marketing" --
  tag "other=accounting"
  ```

  Tenga en cuenta que cuando se agregan etiquetas mediante la AWS CLI, el comando no produce ningún resultado. Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Visualización de las etiquetas de un clúster de Amazon EMR

Si desea ver todas las etiquetas asociadas a un clúster, puede verlas en la consola o en la AWS CLI.

Console

Para ver las etiquetas en un clúster mediante la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. Para ver todas las etiquetas, seleccione la pestaña Etiquetas en la página de detalles del clúster.

AWS CLI

Para ver las etiquetas de un clúster con AWS CLI

Para ver las etiquetas de un clúster mediante el AWS CLI, escriba el `describe-cluster` subcomando con el `--query` parámetro.

- Para ver las etiquetas de un clúster, escribe el siguiente comando y *`j-KT4XXXXXXXX1NM`* sustitúyelo por el ID del clúster.

```
aws emr describe-cluster --cluster-id j-KT4XXXXXX1NM --query Cluster.Tags
```

La salida muestra toda la información de etiquetas sobre el clúster similar a la siguiente:

```
Value: accounting     Value: marketing
Key: other            Key: costCenter
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Eliminación de las etiquetas de un clúster de Amazon EMR

Si ya no necesita una etiqueta, puede eliminarla del clúster.

Console

Para eliminar etiquetas de un clúster con la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. En la pestaña Etiquetas de la página de detalles del clúster, seleccione Administrar etiquetas.
4. Seleccione Eliminar en cada par de clave-valor que desee eliminar.
5. Seleccione Save changes (Guardar cambios).

AWS CLI

Para eliminar las etiquetas de un clúster con AWS CLI

Escriba el subcomando `remove-tags` con el parámetro `--tag-keys`. Al eliminar una etiqueta, solo se necesita el nombre de clave.

- Para eliminar una etiqueta de un clúster, escribe el siguiente comando y *`j-KT4XXXXXXXX1NM`* sustitúyela por el ID del clúster.

```
aws emr remove-tags --resource-id j-KT4XXXXXX1NM --tag-keys "costCenter"
```

> **Note**
> Actualmente no puede eliminar varias etiquetas con un único comando.

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Integración de controladores y aplicaciones de terceros en Amazon EMR

Puede ejecutar varias aplicaciones de macrodatos populares en Amazon EMR con los precios de servicios públicos. Esto significa que paga una tarifa nominal adicional por hora para la aplicación de terceros, mientras el clúster está en ejecución. Le permite usar la aplicación sin tener que adquirir una licencia anual. En las siguientes secciones se describen algunas de las herramientas que puede utilizar con EMR.

Temas

- Utilizar herramientas de inteligencia empresarial con Amazon EMR

Utilizar herramientas de inteligencia empresarial con Amazon EMR

Puede usar herramientas de inteligencia empresarial populares MicroStrategy, como Microsoft Excel y Tableau con Amazon EMRQlikView, para explorar y visualizar sus datos. Muchas de estas herramientas requieren un controlador ODBC (Open Database Connectivity) o JDBC (Java Database Connectivity). Para descargar e instalar los controladores más recientes, consulte http://awssupportdatasvcs.com/bootstrap-actions/Simba/latest/.

Para buscar versiones anteriores de los controladores, consulte http://awssupportdatasvcs.com/bootstrap-actions/Simba/.

Seguridad en Amazon EMR

La seguridad y el cumplimiento son una responsabilidad que usted comparte. AWS Este modelo de responsabilidad compartida puede ayudar a aliviar la carga operativa al AWS operar, administrar y controlar los componentes desde el sistema operativo anfitrión y la capa de virtualización hasta la seguridad física de las instalaciones en las que operan los clústeres de EMR. Usted asume la responsabilidad de gestionar y actualizar los clústeres de Amazon EMR, así como de configurar el software de la aplicación y los controles de seguridad AWS proporcionados. Esta diferenciación de responsabilidad suele denominarse seguridad de la nube en comparación con seguridad en la nube.

- Seguridad de la nube: AWS es responsable de proteger la infraestructura Servicios de AWS en AWS la que se ejecuta. AWS también le proporciona servicios que puede utilizar de forma segura. Auditores independientes prueban y verifican periódicamente la eficacia de nuestra seguridad en el marco de los programas de conformidad de AWS. Para obtener más información sobre los programas de conformidad que se aplican a Amazon EMR, consulte Servicios de AWS en el ámbito del programa de conformidad.
- Seguridad en la nube: también es responsable de realizar todas las tareas de configuración y administración de la seguridad necesarias para proteger un clúster de Amazon EMR. Los clientes que despliegan un clúster de Amazon EMR son responsables de la administración del software de aplicación instalado en las instancias y de la configuración de las funciones AWS proporcionadas, como los grupos de seguridad, el cifrado y el control de acceso, de acuerdo con sus requisitos y las leyes y normativas aplicables.

Esta documentación le permite comprender cómo aplicar el modelo de responsabilidad compartida cuando se utiliza Amazon EMR. Los temas de este capítulo muestran cómo configurar Amazon EMR y utilizar otros Servicios de AWS para cumplir sus objetivos de seguridad y conformidad.

Seguridad de la red y la infraestructura

Como servicio gestionado, Amazon EMR está protegido por los procedimientos de seguridad de la red AWS global que se describen en el documento técnico Amazon Web Services: descripción general de los procesos de seguridad. AWS Los servicios de protección de redes e infraestructuras le ofrecen protecciones detalladas tanto a nivel de host como de red. Amazon EMR admite Servicios de AWS y funciones de la aplicación que abordan los requisitos de conformidad y protección de la red.

- Los grupos EC2 de seguridad de Amazon actúan como un firewall virtual para las instancias de clúster de Amazon EMR, lo que limita el tráfico de red entrante y saliente. Para obtener más información, consulte Control del tráfico de red con grupos de seguridad.
- El acceso público de bloqueo de Amazon EMR (BPA) le impide lanzar un clúster en una subred pública si el clúster tiene una configuración de seguridad que permite el tráfico entrante desde direcciones IP públicas en un puerto. Para obtener más información, consulte Uso del acceso público de bloqueo de Amazon EMR.
- Secure Shell (SSH) le ayuda a proporcionar una forma segura de que los usuarios se conecten a la línea de comandos en instancias de clúster. También puede usar SSH para ver las interfaces web que las aplicaciones alojan en el nodo maestro de un clúster. Para obtener más información, consulte Usar un EC2 key pair para las credenciales SSH y Connect to a cluster.

Actualizaciones de la AMI de Amazon Linux predeterminada para Amazon EMR

> **Important**
>
> Los clústeres de EMR que ejecutan Amazon Linux o Amazon Linux 2 Amazon Machine Images (AMIs) utilizan el comportamiento predeterminado de Amazon Linux y no descargan e instalan automáticamente actualizaciones importantes y críticas del núcleo que requieren un reinicio. Este comportamiento es el mismo que el de otras EC2 instancias de Amazon que ejecutan la AMI de Amazon Linux predeterminada. Si aparecen nuevas actualizaciones de software de Amazon Linux que requieren un reinicio (por ejemplo, actualizaciones del kernel, NVIDIA y CUDA) tras el lanzamiento de una versión de Amazon EMR, las instancias de clúster de Amazon EMR que ejecutan la AMI predeterminada no descargan ni instalan automáticamente esas actualizaciones. Para obtener actualizaciones del kernel, puede personalizar la AMI de Amazon EMR para que utilice la AMI de Amazon Linux más reciente.

En función del nivel de seguridad de la aplicación y del tiempo que lleve ejecutándose un clúster, puede optar por reiniciar el clúster periódicamente para aplicar las actualizaciones de seguridad, o crear una acción de arranque para personalizar la instalación y las actualizaciones de los paquetes. También puede optar por probar y, a continuación, instalar determinadas actualizaciones de seguridad en las instancias del clúster en ejecución. Para obtener más información, consulte Uso de la AMI de Amazon Linux predeterminada para Amazon EMR. Tenga en cuenta que la configuración

de red debe permitir la salida de HTTP y HTTPS a los repositorios de Linux en Amazon S3; de lo contrario, las actualizaciones de seguridad no se realizarán correctamente.

AWS Identity and Access Management con Amazon EMR

AWS Identity and Access Management (IAM) es un AWS servicio que ayuda a un administrador a controlar de forma segura el acceso a los AWS recursos. Los administradores de IAM controlan quién se puede autenticar (iniciar sesión) y autorizar (tener permisos) para utilizar los recursos de Amazon EMR. Las identidades de IAM incluyen usuarios, grupos y roles. El rol de IAM es similar a la de un usuario de IAM, pero no está asociada a una determinada persona y está destinada a ser asumida por cualquier usuario que necesite permisos. Para obtener más información, consulte AWS Identity and Access Management para Amazon EMR. Amazon EMR utiliza varios roles de IAM para ayudarle a implementar controles de acceso para los clústeres de Amazon EMR. La IAM es un AWS servicio que puede utilizar sin coste adicional.

- Función de IAM para Amazon EMR (función EMR): controla la forma en que el servicio Amazon EMR puede acceder a Servicios de AWS otros en su nombre, como el aprovisionamiento de EC2 instancias de Amazon cuando se lanza el clúster de Amazon EMR. Para obtener más información, consulte Configurar las funciones de servicio de IAM para los permisos Servicios de AWS y los recursos de Amazon EMR.
- Función de IAM para EC2 instancias de clúster (perfil de EC2 instancia): función que se asigna a todas las EC2 instancias del clúster de Amazon EMR cuando se lanza la instancia. Los procesos de aplicación que se ejecutan en el clúster utilizan este rol para interactuar con otros Servicios de AWS, como Amazon S3. Para obtener más información, consulte la función de IAM para las EC2 instancias del clúster.
- Rol de IAM para aplicaciones (rol de tiempo de ejecución): un rol de IAM que puede especificar al enviar un trabajo o una consulta a un clúster de Amazon EMR. El trabajo o la consulta que envíe a su clúster de Amazon EMR utiliza el rol de tiempo de ejecución para acceder a AWS los recursos, como los objetos de Amazon S3. Puede especificar roles en tiempo de ejecución con Amazon EMR para los trabajos de Spark y Hive. Al usar roles de tiempo de ejecución, puede aislar los trabajos que se ejecutan en el mismo clúster utilizando diferentes roles de IAM. Para obtener información, consulte Uso de un rol de IAM como rol de tiempo de ejecución con Amazon EMR.

Las identidades de la fuerza laboral se refieren a los usuarios que crean u operan cargas de trabajo en ellas. AWS Amazon EMR proporciona compatibilidad de las identidades del personal con lo siguiente:

- AWS El centro de identidad de IAM (Idc) es el recomendado Servicio de AWS para administrar el acceso de los usuarios a los recursos. AWS Es un lugar único donde puede asignar las identidades de sus empleados y acceder de forma uniforme a varias AWS cuentas y aplicaciones. Amazon EMR respalda las identidades del personal mediante una propagación de identidades fiable. Con una capacidad fiable de propagación de identidades, un usuario puede iniciar sesión en la aplicación y esa aplicación puede transmitir la identidad del usuario a otro usuario Servicios de AWS para autorizar el acceso a los datos o los recursos. Para obtener más información, consulte Habilitar la compatibilidad para el AWS IAM Identity Center con Amazon EMR.

  El Protocolo ligero de acceso a directorios (LDAP) es un protocolo de aplicación estándar del sector, abierto, independiente del proveedor y que permite acceder a la información sobre los usuarios, los sistemas, los servicios y las aplicaciones y mantenerla a través de la red. El LDAP se utiliza habitualmente para la autenticación de usuarios en servidores de identidad como Active Directory (AD) y OpenLDAP. Al habilitar LDAP con clústeres de EMR, usted permite a los usuarios utilizar sus credenciales existentes para autenticarse y acceder a los clústeres. Para obtener más información, consulte habilitación de la compatibilidad para el LDAP con Amazon EMR.

  Kerberos es un protocolo de autenticación de red diseñado para proporcionar una autenticación sólida a client/server las aplicaciones mediante criptografía de clave secreta. Si utiliza Kerberos, Amazon EMR configura Kerberos para las aplicaciones, componentes y subsistemas que instala en el clúster, de forma que se autentiquen entre sí. Para acceder a un clúster con Kerberos configurado, debe haber una entidad principal de Kerberos en el Controlador de Dominio de Kerberos (KDC). Para obtener más información, consulte habilitación de la compatibilidad para Kerberos con Amazon EMR.

Clústeres de un solo inquilino y de varios inquilinos

De forma predeterminada, un clúster está configurado para un solo arrendamiento con el perfil de EC2 instancia como identidad de IAM. En un clúster de un solo inquilino, cada trabajo tiene acceso total y completo al clúster y el acceso a todos los Servicios de AWS recursos se realiza en función del perfil de la EC2 instancia. En un clúster con varios inquilinos, los inquilinos están aislados unos de otros y no tienen acceso total ni completo a los clústeres ni a las EC2 instancias del clúster. La identidad de los clústeres de varios inquilinos son los roles de tiempo de ejecución o los que identifica el personal. En un clúster multiinquilino, también puedes habilitar la compatibilidad con el control de acceso detallado (FGAC) mediante Apache Ranger. AWS Lake Formation Si un clúster tiene habilitadas las funciones de tiempo de ejecución o el FGAC, el acceso al perfil de instancia también está inhabilitado a través de iptables. EC2

> **Important**
>
> Cualquier usuario que tenga acceso a un clúster de un solo inquilino puede instalar cualquier software en el sistema operativo (SO) Linux, cambiar o eliminar los componentes de software instalados por Amazon EMR y afectar a EC2 las instancias que forman parte del clúster. Si quiere asegurarse de que los usuarios no puedan instalar o cambiar las configuraciones de un clúster de Amazon EMR, le recomendamos que habilite la multitenencia para el clúster. Para habilitar la multitenencia en un clúster, habilite la compatibilidad con el rol de tiempo de ejecución, el centro de identidad de AWS IAM, Kerberos o LDAP.

Protección de los datos

Con él AWS, puede controlar sus datos mediante el uso Servicios de AWS de herramientas para determinar cómo están protegidos los datos y quién tiene acceso a ellos. Los servicios como AWS Identity and Access Management (IAM) le permiten administrar de forma segura el acceso Servicios de AWS y los recursos. AWS CloudTrail permite la detección y la auditoría. Amazon EMR le facilita el cifrado de los datos en reposo en Amazon S3 mediante claves gestionadas por usted AWS o totalmente gestionadas por usted. Amazon EMR también admite la activación del cifrado de los datos en tránsito. Para obtener más información, consulte cifrado de datos en reposo y en tránsito.

Control de acceso a los datos

Con el control de acceso a los datos, puede controlar a qué datos puede acceder una identidad de IAM o un identidad de personal. Amazon EMR admite los siguientes controles de acceso:

- Políticas de IAM basadas en la identidad: administre los permisos para los roles de IAM que utilice con Amazon EMR. Las políticas de IAM se pueden combinar con el etiquetado para controlar el acceso de forma individual. cluster-by-cluster Para obtener más información, consulte AWS Identity and Access Management para Amazon EMR.
- AWS Lake Formation centraliza la administración de permisos de sus datos y facilita su uso compartido en toda la organización y de forma externa. Puede usar Lake Formation para permitir un acceso detallado a nivel de columnas a las bases de datos y tablas del catálogo de datos de Glue. AWS Para obtener más información, consulte Uso AWS Lake Formation con Amazon EMR.
- El acceso a Amazon S3 otorga identidades de mapas que mapean identidades en directorios como Active Directory o entidades principales AWS Identity and Access Management (IAM) a conjuntos de datos de S3. Además, el acceso a S3 concede la identidad del usuario final de

registro y la aplicación utilizada para acceder a los datos de S3 en AWS CloudTrail. Para obtener más información, consulte Uso de concesiones de acceso a Amazon S3 con Amazon EMR.

- Apache Ranger es un marco para habilitar, monitorizar y administrar la seguridad integral de los datos en toda la plataforma Hadoop. Amazon EMR admite un control de acceso detallado basado en Apache Ranger para Apache Hive Metastore y Amazon S3. Para más información, consulte Integrar Apache Ranger con Amazon EMR.

Uso de configuraciones de seguridad para configurar la seguridad del clúster de Amazon EMR

Puede utilizar configuraciones de seguridad de Amazon EMR para establecer el cifrado de datos, la autenticación de Kerberos y la autorización de Amazon S3 para EMRFS en los clústeres. Primero, necesita crear una configuración de seguridad. A continuación, la configuración de seguridad estará disponible para su uso y reutilización al crear clústeres.

Puede usar el AWS Management Console, el AWS Command Line Interface (AWS CLI) o el AWS SDKs para crear configuraciones de seguridad. También puede usar una AWS CloudFormation plantilla para crear una configuración de seguridad. Para obtener más información, consulte la Guía AWS CloudFormation del usuario y la referencia de la plantilla para AWS::EMR::SecurityConfiguration.

Temas

- Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI
- Especificar una configuración de seguridad para un clúster de Amazon EMR

Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI

En este tema se describen los procedimientos generales para crear una configuración de seguridad con la consola de Amazon EMR y la AWS CLI, seguido de una referencia sobre los parámetros que comprenden las funciones de cifrado, autenticación e IAM para EMRFS. Para obtener más información sobre estas características, consulte los siguientes temas:

- Cifrar datos en reposo y en tránsito con Amazon EMR
- Uso de Kerberos para la autenticación con Amazon EMR

- Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3

Creación de una configuración de seguridad utilizando la consola

1. Abra la consola Amazon EMR en https://console.aws.amazon.com /emr.
2. En el panel de navegación, elija Security Configurations (Configuraciones de seguridad), Create security configuration (Crear configuración de seguridad).
3. Escriba un nombre para la configuración de seguridad en el campo Name (Nombre).
4. Elija opciones de Cifrado y Autenticación, tal y como se describe en las secciones siguientes y, a continuación, elija Crear.

Para crear una configuración de seguridad mediante el AWS CLI

- Utilice el comando `create-security-configuration`, tal y como se muestra en el ejemplo siguiente.
  - Para*`SecConfigName`*, especifique el nombre de la configuración de seguridad. Este es el nombre que especifica cuando crea un clúster que usa esta configuración de seguridad.
  - En *`SecConfigDef`*, especifique una estructura JSON en línea o la ruta a un archivo JSON local, como por ejemplo, *`file://MySecConfig.json`*. Los parámetros de JSON definen opciones de Cifrado, Roles de IAM para el acceso de EMRFS a Amazon S3 y Autenticación, tal como se describe en las secciones siguientes.

```
aws emr create-security-configuration --name "SecConfigName" --security-
configuration SecConfigDef
```

Configuración del cifrado de datos

Antes de configurar el cifrado en una configuración de seguridad, cree las claves y los certificados que se utilizan para el cifrado. Para obtener más información, consulte Proporcionar claves para cifrado de datos en reposo y Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR.

Al crear una configuración de seguridad, debe especificar dos conjuntos de opciones de cifrado: cifrado de datos en reposo y cifrado de datos en tránsito. Las opciones para el cifrado de datos en

reposo incluyen tanto Amazon S3 con EMRFS y el cifrado de disco local. Las opciones de cifrado en tránsito habilitan las características de cifrado de código abierto para determinadas aplicaciones que admiten Transport Layer Security (TLS). Las opciones en reposo y en tránsito se pueden habilitar juntas o por separado. Para obtener más información, consulte Cifrar datos en reposo y en tránsito con Amazon EMR.

> **Note**
>
> Al utilizarlas AWS KMS, se cobran cargos por el almacenamiento y el uso de las claves de cifrado. Para obtener más información, consulte AWS KMS Precios.

Especificación de las opciones de cifrado con la consola

Elija opciones en Encryption (Cifrado) de acuerdo con las siguientes directrices.

- Elija las opciones en At rest encryption (Cifrado en reposo) para cifrar los datos almacenados en el sistema de archivos.

  Puede elegir cifrar datos en Amazon S3, discos locales o ambas opciones.
- En Cifrado de datos de S3, para Modo de cifrado, seleccione un valor para determinar cómo Amazon EMR cifra los datos de Amazon S3 con EMRFS.

  Lo que haga a continuación depende del modo de cifrado que haya elegido:
  - SSE-S3

    Especifica el cifrado del servidor con claves de cifrado administradas por Amazon S3. No necesita hacer nada más, ya que Amazon S3 se encarga de gestionar las claves.
  - SSE-KMS o CSE-KMS

    Especifica el cifrado del lado del servidor con claves AWS KMS administradas (SSE-KMS) o el cifrado del lado del cliente con claves administradas (CSE-KMS). AWS KMS En AWS KMS key, seleccione una clave. La clave debe existir en la misma región que su clúster de EMR. Para los requisitos de clave, consulte Utilización AWS KMS keys para el cifrado.
  - CSE-Custom

    Especifica el cifrado del cliente utilizando una clave raíz del cliente (CSE-Custom). En Objeto de S3, especifique la ubicación en Amazon S3 o el ARN de Amazon S3 del archivo JAR del proveedor de claves personalizadas. A continuación, en Clase de proveedor de claves,

introduzca el nombre completo de la clase declarada en la aplicación que implementa la interfaz. EncryptionMaterialsProvider

- En Local disk encryption (Cifrado de disco local), seleccione un valor para Key provider type (Tipo de proveedor de clave).
  - AWS KMS key

    Seleccione esta opción para especificar una AWS KMS key. En AWS KMS key, seleccione una clave. La clave debe existir en la misma región que su clúster de EMR. Para obtener más información sobre los requisitos de claves, consulte Utilización AWS KMS keys para el cifrado.

    Cifrado de EBS

    Si lo especifica AWS KMS como proveedor de claves, puede habilitar el cifrado de EBS para cifrar los volúmenes de almacenamiento y los dispositivos raíz de EBS. Para habilitar esta opción, debe conceder el rol de servicio de Amazon EMR `EMR_DefaultRole` con permisos para utilizar la AWS KMS key que especifique. Para obtener más información sobre los requisitos de claves, consulte Habilitación del cifrado de EBS proporcionando permisos adicionales para las claves de KMS.
  - Personalizada

    Seleccione esta opción para especificar un proveedor de claves personalizadas. En Objeto de S3, especifique la ubicación en Amazon S3 o el ARN de Amazon S3 del archivo JAR del proveedor de claves personalizadas. En Clase de proveedor de claves, introduzca el nombre completo de la clase declarada en la aplicación que implementa la interfaz. EncryptionMaterialsProvider El nombre de clase que proporcione aquí debe ser distinto al nombre de clase proporcionado para CSE-Custom.
- Seleccione In-transit encryption (Cifrado en tránsito.) para habilitar las características de cifrado TLS de código abierto para los datos en tránsito. Elija un Certificate provider type (Tipo de proveedor de certificados) de acuerdo con las directrices siguientes:
  - PEM

    Seleccione esta opción para utilizar los archivos PEM que proporcione dentro de un archivo zip. Se requieren dos artefactos dentro del archivo zip: privateKey.pem y certificateChain.pem. Un tercer archivo, trustedCertificates.pem, es opcional. Para obtener más información, consulte Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR. En Objeto de S3, especifique la ubicación en Amazon S3 o el ARN de Amazon S3 del campo del archivo ZIP.

- Personalizada

  Seleccione esta opción para especificar un proveedor de certificados personalizado y, a continuación, en Objeto de S3, especifique la ubicación en Amazon S3 o el ARN de Amazon S3 del archivo JAR del proveedor de certificados personalizado. En Clase de proveedor de claves, introduzca el nombre completo de la clase declarada en la aplicación que implementa la interfaz TLSArtifacts de proveedor.

Especificar las opciones de cifrado mediante el AWS CLI

Las secciones que aparecen a continuación utilizan escenarios de ejemplo para ilustrar una estructura JSON --security-configuration con el formato correcto para diferentes configuraciones y proveedores de claves, así como una referencia para los parámetros JSON y sus valores adecuados.

Opciones de cifrado de datos en tránsito de ejemplo

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está habilitado y el cifrado de datos en reposo está deshabilitado.
- Un archivo ZIP con certificados en Amazon S3 se utiliza como proveedor de claves (consulte Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR para conocer los requisitos de los certificados).

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
 "EncryptionConfiguration": {
  "EnableInTransitEncryption": true,
  "EnableAtRestEncryption": false,
  "InTransitEncryptionConfiguration": {
   "TLSCertificateConfiguration": {
    "CertificateProviderType": "PEM",
    "S3Object": "s3://MyConfigStore/artifacts/MyCerts.zip"
   }
  }
 }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está habilitado y el cifrado de datos en reposo está deshabilitado.
- Se utiliza un proveedor de claves personalizadas (consulte Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR para requisitos de certificado).

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
 "EncryptionConfiguration": {
  "EnableInTransitEncryption": true,
  "EnableAtRestEncryption": false,
  "InTransitEncryptionConfiguration": {
   "TLSCertificateConfiguration": {
    "CertificateProviderType": "Custom",
    "S3Object": "s3://MyConfig/artifacts/MyCerts.jar",
    "CertificateProviderClass": "com.mycompany.MyCertProvider"
   }
  }
  }
}'
```

Opciones de cifrado de datos en reposo de ejemplo

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está deshabilitado y el cifrado de datos en reposo está habilitado.
- SSE-S3 se utiliza para el cifrado de Amazon S3.
- El cifrado del disco local AWS KMS se utiliza como proveedor de claves.

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
 "EncryptionConfiguration": {
  "EnableInTransitEncryption": false,
  "EnableAtRestEncryption": true,
  "AtRestEncryptionConfiguration": {
   "S3EncryptionConfiguration": {
    "EncryptionMode": "SSE-S3"
   },
   "LocalDiskEncryptionConfiguration": {
    "EncryptionKeyProviderType": "AwsKms",
    "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
   }
```

```
  }
  }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está habilitado y hace referencia a un archivo ZIP con certificados PEM en Amazon S3 utilizando el ARN.
- SSE-KMS se utiliza para el cifrado de Amazon S3.
- El cifrado del disco local AWS KMS se utiliza como proveedor de claves.

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
 "EncryptionConfiguration": {
  "EnableInTransitEncryption": true,
  "EnableAtRestEncryption": true,
  "InTransitEncryptionConfiguration": {
   "TLSCertificateConfiguration": {
    "CertificateProviderType": "PEM",
    "S3Object": "arn:aws:s3:::MyConfigStore/artifacts/MyCerts.zip"
   }
  },
  "AtRestEncryptionConfiguration": {
   "S3EncryptionConfiguration": {
    "EncryptionMode": "SSE-KMS",
    "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
   },
   "LocalDiskEncryptionConfiguration": {
    "EncryptionKeyProviderType": "AwsKms",
    "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
   }
  }
 }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está habilitado y hace referencia a un archivo ZIP con certificados PEM en Amazon S3.

- CSE-KMS se utiliza para el cifrado de Amazon S3.
- El cifrado en disco local utiliza un proveedor de claves personalizadas al que hace referencia su ARN.

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
 "EncryptionConfiguration": {
  "EnableInTransitEncryption": true,
  "EnableAtRestEncryption": true,
  "InTransitEncryptionConfiguration": {
   "TLSCertificateConfiguration": {
    "CertificateProviderType": "PEM",
    "S3Object": "s3://MyConfigStore/artifacts/MyCerts.zip"
   }
  },
  "AtRestEncryptionConfiguration": {
   "S3EncryptionConfiguration": {
    "EncryptionMode": "CSE-KMS",
    "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
   },
   "LocalDiskEncryptionConfiguration": {
    "EncryptionKeyProviderType": "Custom",
    "S3Object": "arn:aws:s3:::artifacts/MyKeyProvider.jar",
    "EncryptionKeyProviderClass": "com.mycompany.MyKeyProvider"
   }
  }
 }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está habilitado con un proveedor de claves personalizadas.
- CSE-Custom se utiliza para los datos de Amazon S3.
- El cifrado de disco local utiliza un proveedor de claves personalizadas.

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
 "EncryptionConfiguration": {
```

```
  "EnableInTransitEncryption": "true",
  "EnableAtRestEncryption": "true",
  "InTransitEncryptionConfiguration": {
   "TLSCertificateConfiguration": {
    "CertificateProviderType": "Custom",
    "S3Object": "s3://MyConfig/artifacts/MyCerts.jar",
    "CertificateProviderClass": "com.mycompany.MyCertProvider"
   }
  },
  "AtRestEncryptionConfiguration": {
   "S3EncryptionConfiguration": {
    "EncryptionMode": "CSE-Custom",
    "S3Object": "s3://MyConfig/artifacts/MyCerts.jar",
    "EncryptionKeyProviderClass": "com.mycompany.MyKeyProvider"
    },
   "LocalDiskEncryptionConfiguration": {
    "EncryptionKeyProviderType": "Custom",
    "S3Object": "s3://MyConfig/artifacts/MyCerts.jar",
    "EncryptionKeyProviderClass": "com.mycompany.MyKeyProvider"
   }
  }
 }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está deshabilitado y el cifrado de datos en reposo está habilitado.
- El cifrado de Amazon S3 está habilitado con SSE-KMS.
- Se utilizan varias AWS KMS claves, una por cada depósito de S3, y se aplican excepciones de cifrado a estos depósitos de S3 individuales.
- El cifrado de disco local se ha deshabilitado.

```
aws emr create-security-configuration --name "MySecConfig" --security-configuration '{
   "EncryptionConfiguration": {
     "AtRestEncryptionConfiguration": {
             "S3EncryptionConfiguration": {
           "EncryptionMode": "SSE-KMS",
           "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012",
           "Overrides": [
              {
```

```
                "BucketName": "amzn-s3-demo-bucket1",
                "EncryptionMode": "SSE-S3"
              },
              {
                "BucketName": "amzn-s3-demo-bucket2",
                "EncryptionMode": "CSE-KMS",
                "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
              },
              {
                "BucketName": "amzn-s3-demo-bucket3",
                "EncryptionMode": "SSE-KMS",
                "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
              }
            ]
            }
          },
     "EnableInTransitEncryption": false,
      "EnableAtRestEncryption": true
  }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está deshabilitado y el cifrado de datos en reposo está habilitado.
- El cifrado de Amazon S3 se habilita con SSE-S3 y el cifrado de disco local se deshabilita.

```
aws emr create-security-configuration --name "MyS3EncryptionConfig" --security-
configuration '{
    "EncryptionConfiguration": {
        "EnableInTransitEncryption": false,
        "EnableAtRestEncryption": true,
        "AtRestEncryptionConfiguration": {
            "S3EncryptionConfiguration": {
                "EncryptionMode": "SSE-S3"
            }
        }
     }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está deshabilitado y el cifrado de datos en reposo está habilitado.
- El cifrado del disco local está habilitado AWS KMS como proveedor de claves y el cifrado de Amazon S3 está deshabilitado.

```
aws emr create-security-configuration --name "MyLocalDiskEncryptionConfig" --security-
configuration '{
    "EncryptionConfiguration": {
        "EnableInTransitEncryption": false,
        "EnableAtRestEncryption": true,
        "AtRestEncryptionConfiguration": {
            "LocalDiskEncryptionConfiguration": {
                "EncryptionKeyProviderType": "AwsKms",
                "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
            }
        }
     }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

- El cifrado de datos en tránsito está deshabilitado y el cifrado de datos en reposo está habilitado.
- El cifrado del disco local está habilitado AWS KMS como proveedor de claves y el cifrado de Amazon S3 está deshabilitado.
- El cifrado de la EBS está habilitado.

```
aws emr create-security-configuration --name "MyLocalDiskEncryptionConfig" --security-
configuration '{
    "EncryptionConfiguration": {
        "EnableInTransitEncryption": false,
        "EnableAtRestEncryption": true,
        "AtRestEncryptionConfiguration": {
            "LocalDiskEncryptionConfiguration": {
                "EnableEbsEncryption": true,
                "EncryptionKeyProviderType": "AwsKms",
                "AwsKmsKey": "arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
```

```
            }
        }
     }
}'
```

El ejemplo siguiente ilustra el supuesto siguiente:

SSE-EMR-WAL se utiliza para el cifrado EMR WAL

```
aws emr create-security-configuration --name "MySecConfig" \
    --security-configuration '{
        "EncryptionConfiguration": {
            "EMRWALEncryptionConfiguration":{ },
            "EnableInTransitEncryption":false, "EnableAtRestEncryption":false
        }
    }'
```

`EnableInTransitEncryption` y `EnableAtRestEncryption` aún podrían ser verdaderos, si desea habilitar el cifrado relacionado.

El ejemplo siguiente ilustra el supuesto siguiente:

- SSE-KMS-WAL se utiliza para el cifrado EMR WAL
- El cifrado del lado del servidor AWS Key Management Service se utiliza como proveedor de claves

```
aws emr create-security-configuration --name "MySecConfig" \
    --security-configuration '{
        "EncryptionConfiguration": {
            "EMRWALEncryptionConfiguration":{
                "AwsKmsKey":"arn:aws:kms:us-
east-1:123456789012:key/12345678-1234-1234-1234-123456789012"
                },
            "EnableInTransitEncryption":false, "EnableAtRestEncryption":false
        }
    }'
```

`EnableInTransitEncryption` y `EnableAtRestEncryption` aún podrían ser verdaderos, si desea habilitar el cifrado relacionado.

Referencia de JSON para la configuración de cifrado

En la siguiente tabla se muestran los parámetros JSON de configuración de cifrado y se ofrece una descripción de valores aceptados para cada parámetro.

Parámetro	Descripción
`"EnableInTransitEncryption" : true | false`	Especifique `true` para habilitar el cifrado en tránsito y `false` para deshabilitarlo. Si se omite, se supone el valor `false` y el cifrado en tránsito está deshabilitado.
`"EnableAtRestEncryption": true | false`	Especifique `true` para habilitar el cifrado en reposo y `false` para deshabilitarlo. Si se omite, se supone el valor `false` y el cifrado en reposo está deshabilitado.
Parámetros de cifrado en tránsito	
`"InTransitEncryptionConfiguration" :`	Especifica un conjunto de valores que se utiliza para configurar el cifrado en tránsito cuando `EnableInTransitEncryption` es `true`.
`"CertificateProviderType": "PEM" | "Custom"`	Especifica si se deben utilizar certificados `PEM` a los que se hace referencia con un archivo comprimido o un proveedor de certificados `Custom`. Si `PEM` se especifica, `S3Object` debe ser una referencia a la ubicación en Amazon S3 de un archivo zip que contenga los certificados. Si se especifica Custom, `S3Object` debe ser una referencia a la ubicación en Amazon S3 de un archivo JAR, seguida de una `CertificateProviderClass` entrada.
`"S3Object" : " ZipLocation " | "JarLocation "`	Proporciona la ubicación en Amazon S3 a un archivo zip cuando `PEM` se especifica o a un archivo JAR cuando `Custom` se especifica. El formato puede ser una ruta (por ejemplo, `s3://MyConfig/arti`

Parámetro	Descripción
	`facts/CertFiles.zip`) o un ARN (por ejemplo, `arn:aws:s3:::Code/MyCertProvider.jar)` . Si se especific a un archivo zip, debe contener archivos denominados exactamente `privateKey.pem` y `certificateChain.pem` . Un archivo denominado `trustedCertificates.pem` es opcional.
`"CertificateProviderClass" : "MyClassID "`	Necesario solo si `Custom` se especifica para`CertificateProviderType` . *`MyClassID`* especifica un nombre de clase completo declarado en el archivo JAR, que implementa la interfaz del TLSArtifacts proveedor. Por ejemplo, `com.mycompany.MyCertProvider` .
Parámetros de cifrado en reposo	
`"AtRestEncryptionConfiguration" :`	Especifica un conjunto de valores para el cifrado en reposo cuando `EnableAtRestEncryption` está`true`, incluido el cifrado de Amazon S3 y el cifrado de disco local.
Parámetros de cifrado de Amazon S3	
`"S3EncryptionConfiguration" :`	Especifica un conjunto de valores utilizados para el cifrado de Amazon S3 con el sistema de archivos EMR de Amazon (EMRFS).

Parámetro	Descripción
`"EncryptionMode" : "SSE-S3" | "SSE-KMS" | "CSE-KMS" | "CSE-Custom"`	Especifica el tipo de cifrado de Amazon S3 que se va a utilizar. Si `SSE-S3` se especifica, no se requieren más valores de cifrado de Amazon S3. Si `CSE-KMS` se especifica uno `SSE-KMS` o, se debe especificar un AWS KMS key ARN como valor. `AwsKmsKey` Si se especific a `CSE-Custom`, se deben especificar los valores `S3Object` y `EncryptionKeyProviderClass`.
`"AwsKmsKey" : "`*`MyKeyARN`*`"`	Necesario solo cuando se especifica `CSE-KMS` o `SSE-KMS` para `EncryptionMode`. *`MyKeyARN`* debe ser un ARN totalmente especificado para una clave (por ejemplo, `arn:aws:kms:us-east-1:123456789012:key/12345678-1234-1234-1234-123456789012`).
`"S3Object" : "`*`JarLocation`*`"`	Se requiere solo cuando `CSE-Custom` se especifica para`CertificateProviderType`. *`JarLocation`* proporciona la ubicación en Amazon S3 a un archivo JAR. El formato puede ser una ruta (por ejemplo, `s3://MyConfig/artifacts/MyKeyProvider.jar`) o un ARN (por ejemplo, `arn:aws:s3:::Code/MyKeyProvider.jar)`.

Parámetro	Descripción
`"EncryptionKeyProviderClass" : "MyS3KeyClassID "`	Se requiere solo cuando `CSE-Custom` se especifica para `EncryptionMode` . *`MyS3KeyClassID`* especifica el nombre completo de una clase declarada en la aplicación que implementa la EncryptionMaterialsProvider interfaz; por ejemplo, *`com.mycompany.MyS3KeyProvider`* .
Parámetros de cifrado de disco local	
`"LocalDiskEncryptionConfiguration"`	Especifica el proveedor de claves y valores correspondientes que utilizar para el cifrado de disco local.
`"EnableEbsEncryption": true | false`	Especifique si `true` desea habilitar el cifrado EBS. El cifrado de EBS cifra el volumen del dispositivo raíz de EBS y los volúmenes de almacenamiento adjuntos. Para utilizar el cifrado de EBS, debe especificarlo como su. `AwsKms` `EncryptionKeyProviderType`
`"EncryptionKeyProviderType": "AwsKms" | "Custom"`	Especifica el proveedor de claves. Si `AwsKms` se especifica, se debe especificar un ARN de clave KMS como valor. `AwsKmsKey` Si se especifica **`Custom`**, se deben especificar los valores `S3Object` y `EncryptionKeyProviderClass` .
`"AwsKmsKey : " MyKeyARN"`	Se requiere solo cuando `AwsKms` se especifica para `Type`. *`MyKeyARN`* debe ser un ARN completamente especificado para una clave (por ejemplo, `arn:aws:kms:us-east-1:123456789012:key/12345678-1234-1234-1234-456789012123`).

Parámetro	Descripción
"S3Object" : "*JarLocation* "	Se requiere solo cuando CSE-Custom se especifica paraCertificateProvider Type . *JarLocation* proporciona la ubicación en Amazon S3 a un archivo JAR. El formato puede ser una ruta (por ejemplo, s3://MyConfig/artifacts/MyKeyProvider.jar) o un ARN (por ejemplo, arn:aws:s3:::Code/MyKeyProvider.jar) .
"EncryptionKeyProviderClass" : "*MyLocalDiskKeyClassID* "	Se requiere solo cuando Custom se especifica paraType. *MyLocalDiskKeyClassID* especifica el nombre completo de una clase declarada en la aplicación que implementa la EncryptionMaterialsProvider interfaz; por ejemplo,*com.mycompany.MyLocalDiskKeyProvider* .
Parámetros de cifrado EMR WAL	
"EMRWALEncryptionConfiguration"	Especifica el valor del cifrado EMR WAL.
"AwsKmsKey"	Especifica el ID de clave CMK Arn.

Configuración de la autenticación de Kerberos

Una configuración de seguridad con la configuración de Kerberos solo se puede utilizar en un clúster que se crea con los atributos de Kerberos o si se produce un error. Para obtener más información, consulte Uso de Kerberos para la autenticación con Amazon EMR. Kerberos solo está disponible en la versión 5.10.0 y posteriores de Amazon EMR.

Especificación de la configuración de Kerberos con la consola

Elija opciones en Autenticación de Kerberos de acuerdo con las siguientes directrices.

Parámetro		Descripción
Kerberos		Especifica que Kerberos está habilitado para los clústeres que utilizan esta configuración de seguridad . Si un clúster usa esta configuración de seguridad , también debe tener la configuración de Kerberos especificada o se producirá un error.
Proveedor	KDC dedicado del clúster	Especifica que Amazon EMR crea un KDC en el nodo principal de cualquier clúster que utilice esta configura ción de seguridad. Al crear el clúster, debe especific ar el nombre del dominio y la contraseña de administr ador del KDC. Si es necesario, puede hacer referencia a este KDC desde otros clústeres. Cree esos clústeres con una configuración de seguridad diferente, especifique un KDC externo y utilice el nombre de dominio y la contraseña de administrador del KDC que especifique para el KDC dedicado al clúster.
	KDC externo	Solo está disponible en la versión de Amazon EMR 5.20.0 y posteriores. Especifica que los clústeres que utilizan esta configuración de seguridad autentican las entidades principales de Kerberos mediante un servidor de KDC externo al clúster. No se crea un KDC en el clúster. Al crear el clúster, debe especificar el nombre de dominio y la contraseña de administrador del KDC para el KDC externo.
Ciclo de vida del ticket		Opcional. Especifica el período durante el que un ticket de Kerberos emitido por el KDC es válido en los clústeres que utilizan esta configuración de seguridad. El ciclo de vida de los tickets es limitado por motivos de seguridad. Las aplicaciones de clúster y los servicios renuevan automáticamente los tickets después de que expiren. Los usuarios que se

Parámetro		Descripción
		conecten al clúster mediante SSH utilizando las credenciales de Kerberos tienen que ejecutar el comando `kinit` desde la línea de comandos del nodo principal para la renovación tras la expiración de un ticket.
Relación de confianza entre ámbitos		Especifica una relación de confianza entre ámbitos entre un KDC dedicado a un clúster en los clústeres que utilizan esta configuración de seguridad y un KDC situado en un ámbito de Kerberos diferente. Las entidades principales (normalmente los usuarios) de otro ámbito se autentican en los clústeres que utilizan esta configuración. Se requiere una configura ción adicional en el otro ámbito de Kerberos. Para obtener más información, consulte Tutorial: Configura ción de una relación de confianza entre ámbitos con un dominio de Active Directory.
Propiedad es de la relación de confianza entre ámbitos	Ámbito	Especifica el nombre de ámbito de Kerberos del otro ámbito en la relación de confianza. Por convención, los nombres de ámbito de Kerberos son los mismos que los nombres de dominio, pero en mayúsculas.
	Dominio	Especifica el nombre de dominio del otro ámbito en la relación de confianza.

Parámetro		Descripción
	Servidor de administración	Especifica el nombre de dominio completo (FQDN) o dirección IP del servidor de administración del otro ámbito de la relación de confianza. El servidor de administración y el servidor de KDC suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos para comunicarse. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 749. También se puede especificar el puerto (por ejemplo, `domain.example.com:749`).
	Servidor de KDC	Especifica el nombre de dominio completo (FQDN) o dirección IP del servidor de KDC del otro ámbito de la relación de confianza. El servidor de KDC y el servidor de administración suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 88. También se puede especificar el puerto (por ejemplo, `domain.example.com:88`).
KDC externo		Especifica que el clúster utiliza el KDC externo del clúster.

Parámetro			Descripción
Propiedades de KDC externo	Servidor de administración		Especifica el nombre de dominio completo (FQDN) o la dirección IP del servidor de administración externo. El servidor de administración y el servidor de KDC suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos para comunicarse. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 749. También se puede especificar el puerto (por ejemplo, `domain.example.com:749`).
	Servidor de KDC		Especifica el nombre de dominio completo (FQDN) del servidor de KDC externo. El servidor de KDC y el servidor de administración suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 88. También se puede especificar el puerto (por ejemplo, `domain.example.com:88`).
	Integración de Active Directory		Especifica que la autenticación de la entidad principal de Kerberos está integrada en un dominio de Microsoft Active Directory.
	Propiedades de la integración de Active Directory	Ámbito de Active Directory	Especifica el nombre de ámbito de Kerberos del dominio de Active Directory. Por convención, los nombres de ámbito de Kerberos suelen ser los mismos que los nombres de dominio, pero en mayúsculas.
		Dominio de Active Directory	Especifica el nombre de dominio de Active Directory.

Parámetro			Descripción
		Servidor de Active Directory	Especifica el nombre de dominio completo (FQDN) del controlador de dominio de Microsoft Active Directory.

Especificar la configuración de Kerberos mediante el AWS CLI

En la siguiente tabla de referencia, se muestran los parámetros JSON para la configuración de Kerberos en una configuración de seguridad. Para ver configuraciones de ejemplo, consulte Ejemplos de configuraciones.

Parámetro			Descripción
`"AuthenticationConfiguration": {`			Obligatorio para Kerberos. Especifica que una configuración de autenticación forma parte de esta configuración de seguridad.
	`"KerberosConfiguration": {`		Obligatorio para Kerberos. Especifica las propiedades de configuración de Kerberos.

		`"Provider": "ClusterDedicatedKdc",`
		—o—
		`"Provider: "ExternalKdc",`
			`ClusterDedicatedKdc` especifica que Amazon EMR crea un KDC en el nodo principal de cualquier clúster que utilice esta configuración de seguridad. Al crear el clúster, debe especificar el nombre del dominio y la contraseña de administrador del KDC. Si es necesario, puede hacer referencia a este KDC desde otros clústeres. Cree esos clústeres con una configuración de seguridad diferente, especifique un KDC externo y utilice el nombre de ámbito y la contraseñ

Parámetro	Descripción
	a de administrador del KDC que especificó al crear el clúster con el KDC dedicado a un clúster. *ExternalKdc* especifica que el clúster usa un KDC externo. Amazon EMR no crea un KDC en el nodo principal. Un clúster que utilice esta configuración de seguridad debe especificar el nombre de ámbito y la contraseña de administrador del KDC del KDC externo.
`"ClusterDedicatedKdcConfiguration": {`	Es obligatorio cuando se especifica *ClusterDedicatedKdc* .
`"TicketLifetimeInHours": 24,`	Opcional. Especifica el período durante el que un ticket de Kerberos emitido por el KDC es válido en los clústeres que utilizan esta configuración de seguridad. El ciclo de vida de los tickets es limitado por motivos de seguridad . Las aplicaciones de clúster y los servicios renuevan automáticamente los tickets después de que expiren. Los usuarios que se conecten al clúster mediante SSH utilizando las credenciales de Kerberos tienen que ejecutar el comando `kinit` desde la línea de comandos del nodo principal para la renovación tras la expiración de un ticket.

Parámetro	Descripción
`"CrossRealmTrustConfiguration": {`	Especifica una relación de confianza entre ámbitos entre un KDC dedicado a un clúster en los clústeres que utilizan esta configuración de seguridad y un KDC situado en un ámbito de Kerberos diferente. Las entidades principales (normalmente los usuarios) de otro ámbito se autentican en los clústeres que utilizan esta configuración. Se requiere una configuración adicional en el otro ámbito de Kerberos. Para obtener más información, consulte Tutorial: Configuración de una relación de confianza entre ámbitos con un dominio de Active Directory.
`"Realm": "KDC2.COM",`	Especifica el nombre de ámbito de Kerberos del otro ámbito en la relación de confianza. Por convención, los nombres de ámbito de Kerberos son los mismos que los nombres de dominio, pero en mayúsculas.
`"Domain": "kdc2.com",`	Especifica el nombre de dominio del otro ámbito en la relación de confianza.

Parámetro	Descripción
`"AdminServer": "kdc.com:749 ",`	Especifica el nombre de dominio completo (FQDN) o dirección IP del servidor de administración del otro ámbito de la relación de confianza . El servidor de administración y el servidor de KDC suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos para comunicarse. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 749. También se puede especificar el puerto (por ejemplo, `domain.example.com :749`).
`"KdcServer": "kdc.com:88 "`	Especifica el nombre de dominio completo (FQDN) o dirección IP del servidor de KDC del otro ámbito de la relación de confianza. El servidor de KDC y el servidor de administración suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 88. También se puede especificar el puerto (por ejemplo, `domain.example.com :88`).
`}`	
`}`	

Parámetro	Descripción
"ExternalKdcConfiguration": {	Es obligatorio cuando se especifica `ExternalKdc` .
`"TicketLifetimeInHours": 24,`	Opcional. Especifica el período durante el que un ticket de Kerberos emitido por el KDC es válido en los clústeres que utilizan esta configuración de seguridad. El ciclo de vida de los tickets es limitado por motivos de seguridad . Las aplicaciones de clúster y los servicios renuevan automáticamente los tickets después de que expiren. Los usuarios que se conecten al clúster mediante SSH utilizando las credenciales de Kerberos tienen que ejecutar el comando `kinit` desde la línea de comandos del nodo principal para la renovación tras la expiración de un ticket.
`"KdcServerType": "Single",`	Especifica que se hace referencia a un único servidor de KDC. `Single` es el único valor admitido en este momento.

Parámetro				Descripción
			"AdminServer": "*kdc.com:749* ",	Especifica el nombre de dominio completo (FQDN) o la dirección IP del servidor de administración externo. El servidor de administr ación y el servidor de KDC suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos para comunicarse. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 749. También se puede especificar el puerto (por ejemplo, `domain.example.com` *:749*).
			"KdcServer": "*kdc.com:88* ",	Especifica el nombre de dominio completo (FQDN) del servidor de KDC externo. El servidor de KDC y el servidor de administración suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 88. También se puede especificar el puerto (por ejemplo, `domain.example.com` *:88*).
			`"AdIntegrationConfiguration": {`	Especifica que la autenticación de la entidad principal de Kerberos está integrada en un dominio de Microsoft Active Directory.

Parámetro	Descripción
`"AdRealm": "AD.DOMAIN.COM ",`	Especifica el nombre de ámbito de Kerberos del dominio de Active Directory. Por convención, los nombres de ámbito de Kerberos suelen ser los mismos que los nombres de dominio, pero en mayúsculas.
`"AdDomain": "ad.domain.com "`	Especifica el nombre de dominio de Active Directory.
`"AdServer": "ad.domain.com "`	Especifica el nombre de dominio completo (FQDN) del controlador de dominio de Microsoft Active Directory .
`}`	
`}`	
`}`	
`}`	

Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3

Los roles de IAM para EMRFS le permiten proporcionar diferentes permisos a los datos de EMRFS en Amazon S3. Puede crear asignaciones que especifiquen un rol de IAM que se use para obtener permisos cuando una solicitud de acceso contenga un identificador que haya especificado. El identificador puede ser un usuario o un rol de Hadoop o un prefijo de Amazon S3.

Para obtener más información, consulte Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3.

Especificar las funciones de IAM para EMRFS mediante AWS CLI

A continuación, se muestra un ejemplo de fragmento de JSON para especificar roles de IAM personalizados para EMRFS dentro de una configuración de seguridad. Muestra las asignaciones de roles para los tres tipos de identificadores diferentes, seguidas de una referencia de parámetros.

```
{
  "AuthorizationConfiguration": {
    "EmrFsConfiguration": {
      "RoleMappings": [{
        "Role": "arn:aws:iam::123456789101:role/allow_EMRFS_access_for_user1",
        "IdentifierType": "User",
        "Identifiers": [ "user1" ]
      },{
        "Role": "arn:aws:iam::123456789101:role/allow_EMRFS_access_to_demo_s3_buckets",
        "IdentifierType": "Prefix",
        "Identifiers": [ "s3://amzn-s3-demo-bucket1/","s3://amzn-s3-demo-bucket2/" ]
      },{
        "Role": "arn:aws:iam::123456789101:role/allow_EMRFS_access_for_AdminGroup",
        "IdentifierType": "Group",
        "Identifiers": [ "AdminGroup" ]
      }]
    }
  }
}
```

Parámetro	Descripción
`"AuthorizationConfiguration":`	Obligatorio.
`"EmrFsConfiguration":`	Obligatorio. Contiene asignaciones de roles.
`"RoleMappings":`	Obligatorio. Contiene una o más definiciones de asignación de roles. Las asignaciones de roles se evalúan de forma descendente por orden de aparición. Si una asignación de roles se evalúa como true en una invocació n de datos de EMRFS en Amazon S3, no se evalúan más asignaciones de roles y EMRFS utiliza el rol de IAM especificado para la

Parámetro	Descripción
	solicitud. Las asignaciones de roles tienen los siguientes parámetros obligatorios:
`"Role":`	Especifica el identificador de ARN de un rol de IAM en el formato `arn:aws:iam::` *`account-id`* `:role/`*`role-name`* . Este es el rol de IAM que asume Amazon EMR si la solicitud de EMRFS a Amazon S3 coincide con alguno de los `Identifiers` especificados.
`"IdentifierType":`	Puede ser uno de los siguientes: • `"User"` especifica que los identificadores son uno o más usuarios de Hadoop, que pueden ser usuarios de cuentas de Linux o entidades principales de Kerberos. Cuando la solicitud de EMRFS se origina en el usuario o los usuarios especificados, se asume el rol de IAM. • `"Prefix"` especifica que el identificador es una ubicación de Amazon S3. El rol de IAM se asume para las llamadas a la ubicación o ubicaciones con los prefijos especificados. Por ejemplo, el prefijo `s3://amzn-s3-demo-bucket/` coincide con `s3://amzn-s3-demo-bucket/mydir` y `s3://amzn-s3-demo-bucket/yetanotherdir` . • `"Group"` especifica que los identificadores son uno o más grupos de Hadoop. El rol de IAM se asume si la solicitud proviene de un usuario del grupo o grupos especificados.

Parámetro	Descripción
`"Identifiers":`	Especifica uno o más identificadores del tipo de identificador adecuado. Separe varios identific adores con comas sin espacios.

Configurar las solicitudes de servicios de metadatos a las EC2 instancias de Amazon

Los metadatos de instancia son datos sobre una instancia que se pueden utilizar para configurar o administrar la instancia en ejecución. Para acceder a los metadatos de instancia desde una instancia en ejecución puede utilizar uno de los métodos siguientes:

- Versión 1 (IMDSv1) del servicio de metadatos de instancias: un método de solicitud/respuesta
- Versión 2 (IMDSv2) del servicio de metadatos de instancia: un método orientado a la sesión

Si bien Amazon EC2 admite ambas IMDSv1 opciones IMDSv2, Amazon EMR admite IMDSv2 Amazon EMR 5.23.1, 5.27.1, 5.32 o versiones posteriores y 6.2 o versiones posteriores. En estas versiones, los componentes de Amazon EMR se utilizan IMDSv2 para todas las llamadas del IMDS. Para las llamadas de IMDS en el código de su aplicación, puede utilizar ambos IMDSv1 o configurar el IMDS para que se utilice únicamente IMDSv2 con fines de seguridad adicionales. IMDSv2 Si especificas que IMDSv2 debe usarse, ya IMDSv1 no funciona.

Para obtener más información, consulta Cómo configurar el servicio de metadatos de la instancia en la Guía del EC2 usuario de Amazon.

> **Note**
>
> En versiones anteriores de Amazon EMR 5.x o 6.x, la desactivación provocaba un error en el inicio del clúster, ya que los componentes de Amazon EMR se utilizaban para todas IMDSv1 las llamadas de IMDS. IMDSv1 Al apagarlo IMDSv1, asegúrese de que cualquier software personalizado que utilice esté actualizado a. IMDSv1 IMDSv2

Especificación de la configuración del servicio de metadatos de la instancia mediante AWS CLI

A continuación, se muestra un ejemplo de fragmento de JSON para especificar el servicio de metadatos de EC2 instancias de Amazon (IMDS) dentro de una configuración de seguridad. El uso de una configuración de seguridad personalizada es opcional.

```
{
  "InstanceMetadataServiceConfiguration" : {
      "MinimumInstanceMetadataServiceVersion": integer,
      "HttpPutResponseHopLimit": integer
   }
}
```

Parámetro	Descripción
`"InstanceMetadataServiceConfiguration":`	Si no especifica el IMDS en una configuración de seguridad y utiliza una versión de Amazon EMR que lo requiera IMDSv1, Amazon EMR IMDSv1 utilizará de forma predeterminada la versión mínima del servicio de metadatos de la instancia. Si desea utilizar su propia configuración, se requieren los dos parámetros siguientes.
`"MinimumInstanceMetadataServiceVersion":`	Obligatorio. Indique 1 o 2. Un valor de allows y. 1 IMDSv1 IMDSv2 Un valor de solo 2 permite IMDSv2.
`"HttpPutResponseHopLimit":`	Obligatorio. Límite de saltos de respuesta HTTP PUT deseado para las solicitudes de metadatos de instancia. Cuanto mayor sea el número, más solicitudes de metadatos de instancia pueden viajar. Predeterminado: 1. Puede especificar un número entero del 1 al 64.

Especificación la configuración del servicio de metadatos de la instancia con la consola

Puede configurar el uso del IMDS para un clúster al lanzarlo desde la consola de Amazon EMR.

Para configurar el uso del IMDS mediante la consola:

1. Al crear una nueva configuración de seguridad en la página de configuraciones de seguridad, seleccione Configurar el servicio de metadatos de la EC2 instancia en la configuración del servicio de metadatos de la EC2 instancia. Esta configuración solo se admite en Amazon EMR 5.23.1, 5.27.1, 5.32 o versiones posteriores y 6.2 o versiones posteriores.
2. Para ver la opción Versión mínima del servicio de metadatos de instancias, seleccione una de las siguientes opciones:
   - Desactive IMDSv1 y permita IMDSv2 solo si desea permitir solo IMDSv2 en este clúster. Consulte Transición al uso del servicio de metadatos de instancias, versión 2, en la Guía del EC2 usuario de Amazon.
   - Permita ambas IMDSv1 opciones y IMDSv2 en el clúster, si quiere permitir que IMDSv1 este clúster esté orientado a IMDSv2 las sesiones.
3. Por IMDSv2 ejemplo, también puede configurar el número permitido de saltos de red para el token de metadatos estableciendo el límite de saltos de respuesta HTTP put en un número entero comprendido entre y. 1 64

Para obtener más información, consulta Cómo configurar el servicio de metadatos de la instancia en la Guía del EC2 usuario de Amazon.

Consulte Configurar los detalles de la instancia y Configurar el servicio de metadatos de la instancia en la Guía del EC2 usuario de Amazon.

Especificar una configuración de seguridad para un clúster de Amazon EMR

Puede especificar la configuración de cifrado al crear un clúster especificando la configuración de seguridad. Puede utilizar el AWS Management Console o el AWS CLI.

Console

Para especificar una configuración de seguridad con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Configuración y permisos de seguridad, busque el campo Configuración de seguridad. Seleccione el menú desplegable o elija Examinar para seleccionar el nombre de una configuración de seguridad que haya creado anteriormente. También puede elegir Crear configuración de seguridad para crear una configuración que pueda usar para su clúster.
4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

CLI

Para especificar una configuración de seguridad con el AWS CLI

- Utilice `aws emr create-cluster` para aplicar una configuración de seguridad con `--security-configuration` *`MySecConfig`*, donde *`MySecConfig`* es el nombre de la configuración de seguridad, tal y como se muestra en el siguiente ejemplo. La etiqueta `--release-label` especificada debe ser 4.8.0 o una versión posterior y `--instance-type` puede ser cualquiera disponible.

```
aws emr create-cluster --instance-type m5.xlarge --release-label emr-5.0.0 --
security-configuration mySecConfig
```

Protección de los datos en Amazon EMR

El modelo de responsabilidad AWS compartida se aplica a la protección de datos en Amazon EMR. Como se describe en este modelo, AWS es responsable de proteger la infraestructura global en la que se basa toda la AWS nube. Eres responsable de mantener el control sobre el contenido alojado en esta infraestructura. Este contenido incluye las tareas de configuración y administración de la seguridad AWS que utilice. Para obtener más información sobre la privacidad de datos, consulte Preguntas frecuentes sobre la privacidad de datos. Para obtener información sobre la protección de

datos en Europa, consulta el modelo de responsabilidad compartida de Amazon y la entrada del blog sobre el RGPD en el blog AWS de seguridad.

Para proteger los datos, le recomendamos que proteja las credenciales de las AWS cuentas y configure cuentas individuales con ellas AWS Identity and Access Management. De esta manera, cada usuario recibe únicamente los permisos necesarios para cumplir con sus obligaciones laborales. También recomendamos proteger sus datos de las siguientes maneras:

- Utiliza la autenticación multifactor (MFA) en cada cuenta.
- Use TLS para comunicarse con AWS los recursos. Se requiere usar TLS 1.2.
- Configure la API y el registro de actividad de los usuarios con AWS CloudTrail.
- Utilice soluciones de AWS cifrado, junto con todos los controles de seguridad predeterminados de AWS los servicios.
- Utilice avanzados servicios de seguridad administrados, como Amazon Macie, que lo ayuden a detectar y proteger los datos personales almacenados en Amazon S3.
- Si necesita módulos criptográficos validados FIPS 140-2 al acceder a AWS a través de una interfaz de línea de comandos o una API, utilice un punto de conexión de FIPS. Para obtener más información acerca de los puntos de conexión de FIPS disponibles, consulte Estándar de procesamiento de la información federal (FIPS) 140-2.

Le recomendamos encarecidamente que nunca introduzca información de identificación confidencial, como, por ejemplo, números de cuenta de sus clientes, en los campos de formato libre, como el campo Nombre. Esto incluye cuando trabaja con Amazon EMR u otros AWS servicios mediante la consola, la API o. AWS CLI AWS SDKs Es posible que cualquier dato que ingrese en Amazon EMR u otros servicios se incluya en los registros de diagnóstico. Cuando proporcione una URL a un servidor externo, no incluya información de credenciales en la URL para validar la solicitud para ese servidor.

Cifrar datos en reposo y en tránsito con Amazon EMR

El cifrado de datos ayuda a impedir que los usuarios no autorizados lean los datos en un clúster y sistemas de almacenamiento de datos asociados. Esto incluye los datos guardados en medios persistentes, conocidos como datos en reposo y datos que pueden ser interceptados cuando recorren la red, conocidos como datos en tránsito.

A partir de la versión 4.8.0 de Amazon EMR, puede utilizar configuraciones de seguridad de Amazon EMR para definir configuraciones de cifrado de datos para clústeres de manera más sencilla. Las

configuraciones de seguridad ofrecen ajustes para habilitar la seguridad de los datos en tránsito y de los datos en reposo en volúmenes de Amazon Elastic Block Store (Amazon EBS) y EMRFS en Amazon S3.

Opcionalmente, a partir de la versión 4.1.0 de Amazon EMR y posteriores, puede elegir configurar el cifrado transparente en el HDFS, que no se configura utilizando las configuraciones de seguridad. Para obtener más información, consulte Cifrado transparente en el HDFS en Amazon EMR en la Guía de lanzamiento de Amazon EMR.

Temas

- Opciones de cifrado para Amazon EMR
- Cifrado en reposo mediante una clave KMS del cliente para el servicio EMR WAL
- Creación de claves y certificados para el cifrado de datos con Amazon EMR
- Información sobre el cifrado en tránsito

Opciones de cifrado para Amazon EMR

Con las versiones 4.8.0 y posteriores de Amazon EMR, puede utilizar una configuración de seguridad para especificar las opciones de cifrado de datos en reposo, datos en tránsito o ambos. Al habilitar el cifrado de datos en reposo, puede elegir cifrar datos de EMRFS en Amazon S3, datos en discos locales o ambos. Cada configuración de seguridad que se crea se almacena en Amazon EMR en lugar de en la configuración del clúster, por lo que puede volver a utilizar con facilidad una configuración para especificar ajustes de cifrado de datos cada vez que cree un clúster. Para obtener más información, consulte Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI.

En el siguiente diagrama se muestran las distintas opciones de cifrado de datos disponibles con las configuraciones de seguridad.

Las siguientes opciones de cifrado también están disponibles y no se configuran utilizando una configuración de seguridad:

- Opcionalmente, con las versiones 4.1.0 y posteriores de Amazon EMR, puede elegir configurar el cifrado transparente en el HDFS. Para obtener más información, consulte Cifrado transparente en el HDFS en Amazon EMR en la Guía de lanzamiento de Amazon EMR.

- Si utiliza una versión de Amazon EMR que no admite configuraciones de seguridad, puede configurar manualmente el cifrado para los datos de EMRFS en Amazon S3. Para obtener más información, consulte Especificación del cifrado de Amazon S3 con propiedades de EMRFS.
- Si utiliza una versión de Amazon EMR anterior a 5.24.0, un volumen de dispositivo raíz cifrado de EBS se admite únicamente cuando se utiliza una AMI personalizada. Para obtener más información, consulte Creación de una AMI personalizada con un volumen de dispositivo raíz de Amazon EBS cifrado en la Guía de administración de Amazon EMR.

> **Note**
>
> A partir de la versión 5.24.0 de Amazon EMR, puede utilizar una opción de configuración de seguridad para cifrar los volúmenes de almacenamiento y los dispositivos raíz de EBS si lo especifica como proveedor de claves. AWS KMS Para obtener más información, consulte Cifrado de disco local.

El cifrado de datos requiere las claves y los certificados. Una configuración de seguridad le brinda la flexibilidad de elegir entre varias opciones, incluidas las claves administradas por AWS Key Management Service, las claves administradas por Amazon S3 y las claves y certificados de los proveedores personalizados que usted suministre. Si AWS KMS lo utiliza como proveedor de claves, se aplican cargos por el almacenamiento y el uso de las claves de cifrado. Para obtener más información, consulte Precios de AWS KMS.

Antes de especificar las opciones de cifrado, decida los sistemas de administración de clave y certificado que desee utilizar, para poder crear primero las claves y los certificados o los proveedores personalizados que especifique como parte de la configuración de cifrado.

Cifrado en reposo para datos de EMRFS en Amazon S3

El cifrado de Amazon S3 funciona con objetos del sistema de archivos de Amazon EMR (EMRFS) que se leen y se escriben en Amazon S3. Se especifica el cifrado del servidor (SSE) o el cifrado del cliente (CSE) de Amazon S3 como Modo de cifrado predeterminado al habilitar el cifrado en reposo. También puede especificar métodos de cifrado diferentes para buckets individuales utilizando Per bucket encryption overrides (Reemplazos de cifrado por bucket). Independientemente de si el cifrado de Amazon S3 está habilitado, la seguridad de la capa de transporte (TLS) cifra los objetos de EMRFS en tránsito entre los nodos del clúster de EMR y Amazon S3. Para obtener más información

acerca del cifrado de Amazon S3, consulte Protección de datos mediante cifrado en la Guía del usuario de Amazon Simple Storage Service.

> **Note**
>
> Al utilizarlas AWS KMS, se cobran cargos por el almacenamiento y el uso de las claves de cifrado. Para obtener más información, consulte AWS KMS Precios.

Cifrado del servidor de Amazon S3

Todos los buckets de Amazon S3 tienen el cifrado configurado de forma predeterminada y todos los objetos nuevos que se cargan en un bucket de S3 se cifran automáticamente en reposo. Amazon S3 cifra los datos a nivel de objeto a medida que los escribe en el disco y los descifra cuando se accede a ellos. Para obtener más información sobre SSE, consulte Protección de los datos con el cifrado del servidor en la Guía del usuario de Amazon Simple Storage Service.

Puede elegir entre dos sistemas de administración de claves distintos al especificar SSE en Amazon EMR:

- SSE-S3: Amazon S3 administra las claves en su nombre.
- SSE-KMS: se utiliza una AWS KMS key para configurar políticas adecuadas para Amazon EMR. Para obtener más información sobre los requisitos clave de Amazon EMR, consulte Uso AWS KMS keys para cifrado.

SSE con claves proporcionadas por el cliente (SSE-C) no está disponible para su uso con Amazon EMR.

Cifrado del cliente de Amazon S3

Con el cifrado del cliente de Amazon S3, el proceso de cifrado y descifrado de Amazon S3 se produce en el cliente de EMRFS en su clúster. Los objetos se cifran antes de cargarlos en Amazon S3 y se descifran después de que se descarguen. El proveedor que especifique proporciona la clave de cifrado que utiliza el cliente. El cliente puede usar claves proporcionadas por AWS KMS (CSE-KMS) o una clase de Java personalizada que proporciona la clave raíz del cliente (CSE-C). Los detalles de cifrado son ligeramente diferentes entre CSE-KMS y CSE-C, en función del proveedor especificado y de los metadatos del objeto que se descifra o se cifra. Para obtener más información sobre estas diferencias, consulte Protección de los datos con el cifrado del cliente en la Guía del usuario de Amazon Simple Storage Service.

> **Note**
>
> El CSE de Amazon S3 solo garantiza que los datos de EMRFS intercambiados con Amazon S3 se cifren; no se cifran todos los datos en volúmenes de instancias de clúster. Además, ya que Hue no utiliza EMRFS, los objetos que Hue S3 File Browser escribe en Amazon S3 no se cifran.

Cifrado en reposo para datos en Amazon EMR WAL

Al configurar el cifrado del servidor (SSE) para el registro de escritura anticipada (WAL), Amazon EMR cifra los datos en reposo. Puede elegir a partir de dos sistemas de administración de claves distintos al especificar SSE en Amazon EMR:

SSE-EMR-WAL

Amazon EMR administra las claves por usted. De forma predeterminada, Amazon EMR cifra los datos que ha almacenado en Amazon EMR WAL con SSE-EMR-WAL.

SSE-KMS-WAL

Utiliza una AWS KMS clave para configurar las políticas que se aplican a Amazon EMR WAL. Para obtener más información sobre la configuración del cifrado en reposo para EMR WAL mediante una clave KMS del cliente, consulte Cifrado en reposo mediante una clave KMS del cliente para el servicio EMR WAL.

> **Note**
>
> No puede usar su propia clave con SSE cuando habilita WAL con Amazon EMR. Para obtener más información, consulte Write Ahead Logs (WAL) para Amazon EMR.

Cifrado de disco local

Los siguientes mecanismos funcionan juntos para cifrar discos locales cuando habilita el cifrado de discos locales utilizando una configuración de seguridad de Amazon EMR.

Cifrado HDFS de código abierto

HDFS intercambia datos entre las instancias de clúster durante el procesamiento distribuido. También lee y escribe datos a volúmenes de almacenes de instancias y a los volúmenes de EBS asociado a las instancias. Las siguientes opciones de cifrado de Hadoop de código abierto se activan cuando se habilita el cifrado de disco local:

- Secure Hadoop RPC se define en `Privacy`, que utiliza nivel de seguridad y autenticación simples (SASL).
- Data encryption on HDFS block data transfer se define como `true` y se configura para utilizar el cifrado AES 256.

> **Note**
>
> Puede activar el cifrado de Apache Hadoop adicional habilitando el cifrado en tránsito. Para obtener más información, consulte Cifrado en tránsito. Estos ajustes de cifrado no activan el cifrado transparente de HDFS, que puede configurar manualmente. Para obtener más información, consulte Cifrado transparente en el HDFS en Amazon EMR en la Guía de lanzamiento de Amazon EMR.

Cifrado del almacén de instancias

En el EC2 caso de los tipos que utilizan el volumen del almacén de instancias NVMe basado SSDs , el NVMe cifrado se utiliza independientemente de la configuración de cifrado de Amazon EMR. Para obtener más información, consulta los volúmenes NVMe SSD en la Guía del EC2 usuario de Amazon. En otros volúmenes de almacén de instancias, Amazon EMR utiliza LUKS para cifrar el volumen de almacén de instancias cuando se ha habilitado el cifrado de disco local independientemente de si los volúmenes de EBS se han cifrado utilizando el cifrado de EBS o LUKS.

Cifrado de volumen de EBS

Si crea un clúster en una región en la que el EC2 cifrado de volúmenes de EBS por Amazon está habilitado de forma predeterminada para su cuenta, los volúmenes de EBS se cifran incluso si el cifrado del disco local no está habilitado. Para obtener más información, consulte Cifrado predeterminado en la Guía del EC2 usuario de Amazon. Con el cifrado de disco local activado en una configuración de seguridad, la configuración de Amazon EMR tiene prioridad sobre la configuración de Amazon para las instancias de EC2 encryption-by-default clúster. EC2

Las siguientes opciones están disponibles para volúmenes de cifrado de EBS que utilizan una configuración de seguridad:

- Cifrado EBS: a partir de la versión 5.24.0 de Amazon EMR, puede optar por habilitar el cifrado EBS. La opción de cifrado de EBS cifra el volumen de dispositivo raíz de EBS y los volúmenes de almacenamiento adjuntos. La opción de cifrado de EBS solo está disponible si la especificas AWS Key Management Service como proveedor de claves. Recomendamos el uso del cifrado de EBS.
- Cifrado LUKS: si decide utilizar el cifrado LUKS para los volúmenes de Amazon EBS, el cifrado LUKS se aplica únicamente a los volúmenes de almacenamiento adjuntos, no al volumen del dispositivo raíz. Para obtener más información sobre el cifrado de LUKS, consulte la especificación de LUKS en disco.

  Para su proveedor de claves, puede configurar uno AWS KMS key con políticas adecuadas para Amazon EMR o una clase Java personalizada que proporcione los artefactos de cifrado. Al utilizarlas AWS KMS, se cobran cargos por el almacenamiento y el uso de las claves de cifrado. Para más información, consulte Precios de AWS KMS.

> **Note**
>
> Para comprobar si el cifrado de EBS está habilitado en el clúster, se recomienda utilizar una llamada a la API `DescribeVolumes`. Para obtener más información, consulte DescribeVolumes. La ejecución de `lsblk` en el clúster solo comprobará el estado del cifrado LUKS, en lugar del cifrado de EBS.

Cifrado en tránsito

Hay habilitados diversos mecanismos de cifrado con el cifrado en tránsito. Se trata de características de código abierto, específicas de la aplicación y podrían variar según la versión de Amazon EMR. Para habilitar el cifrado en tránsito, utilice the section called “Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI” en Amazon EMR. Para los clústeres de EMR con el cifrado en tránsito habilitado, Amazon EMR configura automáticamente las configuraciones de las aplicaciones de código abierto para habilitar el cifrado en tránsito. Para los casos de uso avanzados, puede configurar las configuraciones de las aplicaciones de código abierto directamente para anular el comportamiento predeterminado en Amazon EMR. Para obtener más información, consulte la matriz de compatibilidad de cifrado en tránsito y Configuración de aplicaciones.

Consulte lo siguiente para obtener información más específica sobre las aplicaciones de código abierto relacionadas con el cifrado en tránsito:

- Cuando habilita el cifrado en tránsito con una configuración de seguridad, Amazon EMR habilita el cifrado en tránsito para todos los puntos de conexión de aplicaciones de código abierto que admiten el cifrado en tránsito. La compatibilidad con el cifrado en tránsito para los distintos puntos de conexión de las aplicaciones varía según la versión de lanzamiento de Amazon EMR. Para obtener más información, consulte la matriz de compatibilidad de cifrado en tránsito.
- Puede anular las configuraciones de código abierto, lo que le permite hacer lo siguiente:
  - Deshabilite la verificación del nombre de host TLS si los certificados TLS proporcionados por el usuario no cumplen con los requisitos
  - Deshabilite el cifrado en tránsito para determinados puntos de conexión en función de sus requisitos de rendimiento y compatibilidad
  - Controle qué versiones de TLS y conjuntos de cifrado desea utilizar.

  Puede encontrar más detalles sobre las configuraciones específicas de la aplicación en la matriz de compatibilidad de cifrado en tránsito
- Además de habilitar el cifrado en tránsito con una configuración de seguridad, algunos canales de comunicación también requieren configuraciones de seguridad adicionales para poder habilitar el cifrado en tránsito. Por ejemplo, algunos puntos de conexión de aplicaciones de código abierto utilizan la capa de autenticación y seguridad simple (SASL) para el cifrado en tránsito, lo que requiere que la autenticación de Kerberos esté habilitada en la configuración de seguridad del clúster de EMR. Para obtener más información sobre estos puntos de conexión, consulte la matriz de compatibilidad de cifrado en tránsito.
- Se recomienda usar un software que admita TLS v1.2 o posterior. Amazon EMR incluye la distribución JDK de Corretto predeterminada, que determina qué versiones de TLS, conjuntos de cifrado y tamaños de clave permiten las redes de código abierto que se ejecutan en Java. EC2 En este momento, la mayoría de marcos de código abierto aplican TLS v1.2 o posterior para Amazon EMR 7.0.0 y versiones posteriores. Esto se debe a que la mayoría de marcos de código abierto se ejecutan en Java 17 para Amazon EMR 7.0.0 y versiones posteriores. Es posible que las versiones anteriores de Amazon EMR admitan TLS v1.0 y v1.1 porque consumen versiones anteriores de Java, pero Corretto JDK podría cambiar las versiones de TLS que admite Java, lo que podría afectar a las versiones de Amazon EMR existentes.

Tiene que especificar los artefactos de cifrado utilizados para el cifrado en tránsito de una de estas dos maneras: facilitando un archivo comprimido con los certificados que carga en Amazon S3, o

bien, haciendo referencia a una clase de Java personalizada que proporcione artefactos de cifrado. Para obtener más información, consulte Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR.

Cifrado en reposo mediante una clave KMS del cliente para el servicio EMR WAL

Los registros de escritura anticipada (WAL) de EMR proporcionan soporte clave de KMS al cliente. encryption-at-rest Los siguientes detalles detallados sobre cómo se integra Amazon EMR WAL con: AWS KMS

Los registros de escritura anticipada (WAL) de EMR interactúan AWS durante las siguientes operaciones:`CreateWAL`,,,,, `AppendEdit` `ArchiveWALCheckPoint` `CompleteWALFlush` `DeleteWAL` `GetCurrentWALTimeReplayEdits`, de forma `EMR_EC2_DefaultRole` predeterminada. Cuando se `TrimWAL` invoca alguna de las operaciones anteriores de la lista, la WAL de EMR realiza `Decrypt` y contra la clave KMS. `GenerateDataKey`

Consideraciones

Tenga en cuenta lo siguiente cuando utilice el cifrado AWS KMS basado para EMR WAL:

- La configuración de cifrado no se puede cambiar después de crear un EMR WAL.
- Si utiliza el cifrado de KMS con su propia clave de KMS, la clave debe estar en la misma región que su clúster de Amazon EMR.
- Usted es responsable de mantener todos los permisos de IAM necesarios y se recomienda no revocar los permisos necesarios durante la vigencia del WAL. De lo contrario, provocará escenarios de error inesperados, como la imposibilidad de eliminar el EMR WAL, ya que la clave de cifrado asociada no existe.
- El uso de las AWS KMS claves conlleva un coste. Para obtener más información, consulte Precios de AWS Key Management Service.

Permisos de IAM necesarios

Para utilizar la clave KMS de su cliente para cifrar la WAL de EMR en reposo, debe asegurarse de establecer los permisos adecuados para el rol de cliente de EMR WAL y el director de servicio de EMR WAL. `emrwal.amazonaws.com`

Permisos para el rol de cliente EMR WAL

A continuación, se muestra la política de IAM necesaria para el rol de cliente WAL de EMR:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "kms:Decrypt",
        "kms:GenerateDataKey*"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowKMSDecrypt"
    }
  ]
}
```

El cliente EMR WAL del clúster EMR se utilizará de forma predeterminada. `EMR_EC2_DefaultRole` Si utilizas un rol diferente para el perfil de instancia en el clúster de EMR, asegúrate de que cada rol tenga los permisos adecuados.

Para obtener más información sobre la administración de la política de roles, consulta Cómo añadir y eliminar permisos de identidad de IAM.

Permisos para la política de claves de KMS

Debe asignar el rol de cliente EMR WAL y el servicio `Decrypt` y `GenerateDataKey*` permiso EMR WAL en su política de KMS. Para obtener más información sobre la administración de políticas clave, consulte la política clave de KMS.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
```

```
        "kms:Decrypt",
        "kms:GenerateDataKey*"
      ],
      "Resource": [
        "arn:aws:kms:*:123456789012:key/*"
      ],
      "Sid": "AllowKMSDecrypt"
    }
  ]
}
```

El rol especificado en el fragmento puede cambiar si cambias el rol predeterminado.

Supervisión de la interacción de Amazon EMR WAL con AWS KMS

Contexto de cifrado WAL de Amazon EMR

Un contexto de cifrado es un conjunto de pares clave-valor que contiene datos arbitrarios no secretos. Al incluir un contexto de cifrado en una solicitud de cifrado de datos, vincula AWS KMS criptográficamente el contexto de cifrado a los datos cifrados. Para descifrar los datos, es necesario pasar el mismo contexto de cifrado.

En sus solicitudes GenerateDataKeyy en las de Decrypt AWS KMS, Amazon EMR WAL utiliza un contexto de cifrado con un par nombre-valor que identifican el nombre WAL de EMR.

```
"encryptionContext": {
    "aws:emrwal:walname": "111222333444555-testworkspace-emrwalclustertest-
emrwaltestwalname"
}
```

Puede utilizar el contexto de cifrado para identificar estas operaciones criptográficas en los registros y registros de auditoría, como AWS CloudTrail Amazon CloudWatch Logs, y como condición para la autorización en las políticas y las concesiones.

Creación de claves y certificados para el cifrado de datos con Amazon EMR

Antes de especificar las opciones de cifrado mediante una configuración de seguridad, decida el proveedor que desea usar para las claves y los artefactos de cifrado. Por ejemplo, puede usar AWS KMS o un proveedor personalizado que cree. A continuación, cree las claves o el proveedor tal y como se describe en esta sección.

Proporcionar claves para cifrado de datos en reposo

Puede usar AWS Key Management Service (AWS KMS) o un proveedor de claves personalizado para el cifrado de datos en reposo en Amazon EMR. Cuando las usa AWS KMS, se cobran cargos por el almacenamiento y el uso de las claves de cifrado. Para obtener más información, consulte Precios de AWS KMS.

En este tema se ofrecen detalles sobre las políticas de claves para una clave de KMS que se vaya a usar con Amazon EMR, así como directrices y ejemplos de código para escribir una clase de proveedor de claves personalizadas para el cifrado de Amazon S3. Para más información sobre la creación de claves, consulte Creación de claves en la Guía para desarrolladores de AWS Key Management Service .

Utilización AWS KMS keys para el cifrado

La clave de AWS KMS cifrado debe crearse en la misma región que la instancia de clúster de Amazon EMR y los buckets de Amazon S3 que se utilizan con EMRFS. Si la clave que especifica está en una cuenta diferente de la que usa para configurar un clúster, debe especificar la clave mediante su ARN.

El rol del perfil de EC2 instancia de Amazon debe tener permisos para usar la clave de KMS que especifiques. El rol predeterminado del perfil de instancia en Amazon EMR es `EMR_EC2_DefaultRole`. Si usa un rol diferente para el perfil de instancia o usa roles de IAM para las solicitudes de EMRFS a Amazon S3, asegúrese de agregar cada rol como usuario clave, según corresponda. Esto proporciona al rol los permisos para utilizar la clave de KMS. Para obtener más información, consulte Uso de políticas de claves en la Guía para desarrolladores de AWS Key Management Service y Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3.

Puede usarlo AWS Management Console para añadir su perfil de instancia o perfil de EC2 instancia a la lista de usuarios clave de la clave de KMS especificada, o puede utilizar el SDK AWS CLI o un AWS SDK para adjuntar una política de claves adecuada.

Tenga en cuenta que Amazon EMR solo admite claves de KMS simétricas. No se puede utilizar una clave KMS asimétrica para cifrar datos en reposo en un clúster de Amazon EMR. Para obtener ayuda para determinar si una clave de KMS es simétrica o asimétrica, consulte Identificación de claves de KMS simétricas y asimétricas.

El siguiente procedimiento describe cómo agregar el perfil de instancia de Amazon EMR predeterminado, `EMR_EC2_DefaultRole` como un usuario clave con la AWS Management

Console. Se supone que ya ha creado una clave de KMS. Para crear una nueva clave de KMS, consulte Creación de claves en la Guía para desarrolladores de AWS Key Management Service .

Para añadir el perfil de EC2 instancia de Amazon EMR a la lista de usuarios de claves de cifrado

1. Inicie sesión en la consola AWS Key Management Service (AWS KMS) AWS Management Console y ábrala en https://console.aws.amazon.com/kms.
2. Para cambiarla Región de AWS, usa el selector de regiones en la esquina superior derecha de la página.
3. Seleccione el alias de la clave de KMS que desee modificar.
4. En la página de detalles de la clave, en Key Users (Usuarios de claves), seleccione Add (Añadir).
5. En el cuadro de diálogo Add key users (Añadir usuarios clave), seleccione el rol adecuado. El nombre del rol predeterminado es `EMR_EC2_DefaultRole`.
6. Elija Agregar.

Habilitación del cifrado de EBS proporcionando permisos adicionales para las claves de KMS

A partir de la versión 5.24.0 de Amazon EMR, puede cifrar el dispositivo raíz y los volúmenes de almacenamiento de EBS utilizando una opción de configuración de seguridad. Para activar esta opción, debe especificarla AWS KMS como proveedor de claves. Además, debe conceder al rol `EMR_DefaultRole` de servicio permisos para usar el AWS KMS key que especifique.

Puede utilizar la función de servicio AWS Management Console para añadir la función de servicio a la lista de usuarios clave de la clave de KMS especificada, o bien puede utilizar la función de servicio AWS CLI o un AWS SDK para adjuntar una política de claves adecuada.

El siguiente procedimiento describe cómo utilizar el AWS Management Console para añadir el rol de servicio Amazon EMR predeterminado `EMR_DefaultRole` como usuario clave. Se supone que ya ha creado una clave de KMS. Para crear una nueva clave de KMS, consulte Creación de claves en la Guía para desarrolladores de AWS Key Management Service .

Para añadir el rol de servicio de Amazon EMR a la lista de usuarios de claves de cifrado

1. Inicie sesión en la consola AWS Key Management Service (AWS KMS) AWS Management Console y ábrala en https://console.aws.amazon.com/kms.
2. Para cambiarla Región de AWS, usa el selector de regiones en la esquina superior derecha de la página.

3. Elija Claves administradas por el cliente en la barra lateral izquierda.
4. Seleccione el alias de la clave de KMS que desee modificar.
5. En la página de detalles de la clave, en Key Users (Usuarios de claves), seleccione Add (Añadir).
6. En la sección Añadir usuarios de clave, seleccione el rol adecuado. El nombre del rol de servicio predeterminado para Amazon EMR es `EMR_DefaultRole`.
7. Elija Agregar.

Creación de un proveedor de claves personalizadas

Cuando se utiliza una configuración de seguridad, es necesario especificar un nombre de clase de proveedor distinto para el cifrado de disco local y el cifrado de Amazon S3. Los requisitos del proveedor de claves personalizadas dependen de si utiliza el cifrado de disco local y el cifrado de Amazon S3, así como de la versión de lanzamiento de Amazon EMR.

Según el tipo de cifrado que utilice al crear un proveedor de claves personalizado, la aplicación también debe implementar diferentes EncryptionMaterialsProvider interfaces. Ambas interfaces están disponibles en la versión 1.11.0 y posteriores del AWS SDK for Java.

- Para implementar el cifrado de Amazon S3, utilice com.amazonaws.services.s3.model. EncryptionMaterialsProvider interfaz.
- Para implementar el cifrado del disco local, usa com.amazonaws.services.elasticmapreduce.spi.security. EncryptionMaterialsProvider interfaz.

Puede utilizar cualquier estrategia para proporcionar materiales de cifrado para la implementación. Por ejemplo, podría elegir proporcionar materiales de cifrado estáticos o integrar con un sistema de administración de claves más complejo.

Si utiliza el cifrado de Amazon S3, debe utilizar los algoritmos de cifrado AES/GCM/NoPaddingpara los materiales de cifrado personalizados.

Si utiliza el cifrado de disco local, el algoritmo de cifrado que se utilizará para los materiales de cifrado personalizados varía según la versión del EMR. Para Amazon EMR 7.0.0 y versiones anteriores, debe usar. AES/GCM/NoPadding Para Amazon EMR 7.1.0 y versiones posteriores, debe usar AES.

La EncryptionMaterialsProvider clase obtiene los materiales de cifrado por contexto de cifrado. Amazon EMR rellena el contexto de cifrado en tiempo de ejecución para ayudar al intermediario a determinar qué materiales de cifrado debe devolver.

Example Ejemplo: uso de un proveedor de claves de cifrado personalizadas para el cifrado de Amazon S3 con EMRFS

Cuando Amazon EMR obtiene los materiales de cifrado de la EncryptionMaterialsProvider clase para realizar el cifrado, EMRFS rellena opcionalmente el argumento MaterialsDescription con dos campos: el URI de Amazon S3 del objeto y JobFlowId el del clúster, que la clase puede utilizar para devolver materiales de cifrado de forma selectiva. EncryptionMaterialsProvider

Por ejemplo, el proveedor podría devolver claves distintas para diferentes prefijos URI de Amazon S3. Se trata de la descripción de los materiales de cifrado devuelta que se almacena finalmente con el objeto de Amazon S3 en lugar del valor materialsDescription que genera EMRFS y se transfiere al proveedor. Al descifrar un objeto de Amazon S3, la descripción del material de cifrado se pasa a la EncryptionMaterialsProvider clase para que, una vez más, devuelva de forma selectiva la clave correspondiente para descifrar el objeto.

A continuación se EncryptionMaterialsProvider proporciona una implementación de referencia. Otro proveedor personalizado, EMRFSRSAEncryptionMaterialsProvider, está disponible en GitHub.

```
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;
  }
}
```

Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR

Con la versión 4.8.0 o posterior de Amazon EMR, dispone de dos opciones para especificar artefactos para el cifrado de datos en tránsito utilizando una configuración de seguridad:

- Puede crear manualmente certificados PEM, incluirlos en un archivo .zip y, a continuación, hacer referencia al archivo .zip en Amazon S3.
- Puede implementar un proveedor de certificados personalizado como una clase Java. Deberá especificar el archivo JAR de la aplicación en Amazon S3 y, a continuación, proporcionar el nombre de la clase completa del proveedor tal como se declara en la aplicación. La clase debe implementar la interfaz TLSArtifactsde proveedor disponible a partir de la AWS SDK para Java versión 1.11.0.

Amazon EMR descarga automáticamente artefactos en cada nodo del clúster y, posteriormente, los utiliza para implementar las características de cifrado en tránsito de código abierto. Para obtener más información sobre las opciones disponibles, consulte Cifrado en tránsito.

Uso de certificados PEM

Cuando especifique un archivo zip para el cifrado en tránsito, la configuración de seguridad espera que los archivos PEM dentro del archivo zip se nombren exactamente tal y como aparecen a continuación:

Certificados de cifrado en tránsito

Nombre de archivo	Obligatorio/opcional	Detalles
privateKey.pem	Obligatorio	Clave privada
certificateChain.pem	Obligatorio	Cadena de certificados
trustedCertificates.pem	Opcional	Le recomendamos que proporcione un certificado que no esté firmado por la autoridad de certificación (CA) de raíz de confianza predeterminada de Java o una CA intermedia que enlace a la CA de raíz de confianza predeterminada de Java. No le recomendamos que utilice el sistema público CAs cuando utilice certificados comodín o cuando deshabilite la verificación del nombre de host.

Es probable que desee configurar el archivo PEM de clave privada como un certificado comodín que permita el acceso al dominio de Amazon VPC en el que residen las instancias del clúster. Por ejemplo, si el clúster se encuentran en la región us-east-1 (Norte de Virginia), puede elegir

especificar un nombre común en la configuración de certificado que permita el acceso al clúster especificando `CN=*.ec2.internal` en la definición del sujeto del certificado. Si el clúster reside en us-west-2 (Oregón), puede especificar `CN=*.us-west-2.compute.internal`.

Si el archivo PEM proporcionado en el artefacto de cifrado no tiene un carácter comodín para el dominio en el nombre común, debe cambiar el valor de `hadoop.ssl.hostname.verifier` a `ALLOW_ALL`. Para hacerlo en las versiones 7.3.0 y posteriores de Amazon EMR, añada la clasificación `core-site` cuando envíe las configuraciones a un clúster. En las versiones anteriores a la 7.3.0, añada la configuración `"hadoop.ssl.hostname.verifier": "ALLOW_ALL"` directamente al archivo `core-site.xml`. Este cambio es necesario porque el verificador de nombres de host predeterminado requiere un nombre de host sin el comodín porque todos los hosts del clúster lo utilizan. Para obtener más información sobre la configuración del clúster de EMR en una instancia de Amazon VPC, consulte Configuración de redes en una VPC para Amazon EMR.

En el siguiente ejemplo, se muestra cómo utilizar OpenSSL para generar un certificado X.509 autofirmado con una clave privada RSA de 2048 bits. La clave permite el acceso a las instancias de clúster de Amazon EMR en la región `us-west-2` (Oregón), tal y como se especifica mediante el nombre de dominio *`*.us-west-2.compute.internal`* como nombre común.

Se especifican otros elementos del sujeto opcionales como país (C), estado (S) y configuración regional (L). Dado que se genera un certificado autofirmado, el segundo comando del ejemplo copia el archivo `certificateChain.pem` en el archivo `trustedCertificates.pem`. El tercer comando utiliza `zip` para crear el archivo `my-certs.zip` que contiene los certificados.

> **Important**
>
> Este ejemplo es solo una demostración. proof-of-concept El uso de certificados autofirmados no se recomienda y presenta un posible riesgo para la seguridad. En el caso de los sistemas de producción, utilice una autoridad de certificación (CA) de confianza para emitir certificados.

```
$ openssl req -x509 -newkey rsa:2048 -keyout privateKey.pem -out certificateChain.pem
 -days 365 -nodes -subj '/C=US/ST=Washington/L=Seattle/O=MyOrg/OU=MyDept/CN=*.us-
west-2.compute.internal'
$ cp certificateChain.pem trustedCertificates.pem
$ zip -r -X my-certs.zip certificateChain.pem privateKey.pem trustedCertificates.pem
```

Información sobre el cifrado en tránsito

Puede configurar un clúster de EMR para ejecutar marcos de código abierto como Apache Spark, Apache Hive y Presto. Cada uno de estos marcos de código abierto tiene un conjunto de procesos que se ejecutan en las instancias de un clúster. EC2 Cada uno de estos procesos puede alojar puntos de conexión de red para la comunicación de red.

Si el cifrado en tránsito está habilitado en un clúster de EMR, los diferentes puntos de conexión de la red utilizan diferentes mecanismos de cifrado. Consulte las siguientes secciones para obtener más información sobre los puntos de conexión de red específicos del marco de código abierto compatibles con el cifrado en tránsito, los mecanismos de cifrado relacionados y qué versión de Amazon EMR ha añadido esta compatibilidad. Cada aplicación de código abierto también puede tener diferentes prácticas recomendadas y configuraciones de marco de código abierto que puede cambiar.

Para obtener la máxima cobertura de cifrado en tránsito, le recomendamos que habilite tanto el cifrado en tránsito como Kerberos. Si solo habilita el cifrado en tránsito, el cifrado en tránsito solo estará disponible para los puntos de conexión de la red que admiten TLS. Kerberos es necesario porque algunos puntos de conexión de red de código abierto utilizan la capa de autenticación y seguridad simple (SASL) para el cifrado en tránsito.

Tenga en cuenta que no se incluye ningún marco de código abierto que no sea compatible con las versiones 7.x.x de Amazon EMR.

Spark

Al habilitar el cifrado en tránsito en las configuraciones de seguridad, `spark.authenticate` se establece automáticamente en `true` y utiliza el cifrado basado en AES para las conexiones RPC.

A partir de Amazon EMR 7.3.0, si utiliza el cifrado en tránsito y la autenticación Kerberos, no podrá utilizar las aplicaciones de Spark que dependan del metaalmacén de Hive. Hive 3 corrige este problema en HIVE-16340. HIVE-44114 resuelve completamente este problema cuando Spark de código abierto puede actualizarse a Hive 3. Mientras tanto, puede establecer `hive.metastore.use.SSL` en `false` para tratar de solucionar este problema. Para obtener más información, consulte Configuración de aplicaciones.

Para obtener más información, consulte Seguridad de Spark en la documentación de Apache Spark.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Servidor de historial de Spark	spark.ssl.history.port	18480	TLS	emr-5.3.0+, emr-6.0.0+, emr-7.0.0+
Interfaz de usuario de Spark	spark.ui.port	4440	TLS	emr-5.3.0+, emr-6.0.0+, emr-7.0.0+
Controlador de Spark	spark.driver.port	Dinámico	Cifrado basado en AES de Spark	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
Ejecutor de Spark	Puerto ejecutor (sin configuración con nombre)	Dinámico	Cifrado basado en AES de Spark	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
HILO NodeManager	spark.shuffle.service.port[1]	7337	Cifrado basado en AES de Spark	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+

[1] `spark.shuffle.service.port` está alojado en YARN NodeManager pero solo lo usa Apache Spark.

Hadoop YARN

El RPC seguro de Hadoop está configurado para el cifrado en tránsito basado en `privacy` SASL y lo utiliza. Esto requiere que la autenticación de Kerberos esté habilitada en la

configuración de seguridad. Si no desea el cifrado en tránsito para Hadoop RPC, configure `hadoop.rpc.protection = authentication`. Se recomienda usar la configuración predeterminada para garantizar la máxima seguridad.

Si sus certificados TLS no cumplen los requisitos de verificación del nombre de host de TLS, puede configurar `hadoop.ssl.hostname.verifier = ALLOW_ALL`. Se recomienda usar la configuración predeterminada de `hadoop.ssl.hostname.verifier = DEFAULT`, que exige la verificación del nombre de host de TLS.

Para deshabilitar HTTPS en los puntos de conexión de la aplicación web YARN, configure `yarn.http.policy = HTTP_ONLY`. Esto hace que el tráfico a estos puntos de conexión permanezca sin cifrar. Se recomienda usar la configuración predeterminada para garantizar la máxima seguridad.

Para obtener más información, consulte Hadoop in Secure Mode en la documentación de Apache Hadoop:

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
ResourceManager	yarn.resourcemanager.webapp.address	8088	TLS	emr-7.3.0+
ResourceManager	yarn.resourcemanager.resource-tracker.address	8025	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
ResourceManager	yarn.resourcemanager.scheduler.address	8030	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
ResourceManager	yarn.resourcemanager.address	8032	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
ResourceManager	yarn.resourcemanager.admin.address	8033	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
TimelineServer	yarn.timeline-service.address	10200	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
TimelineServer	yarn.timeline-service.webapp.address	8188	TLS	emr-7.3.0+
WebApplicationProxy	yarn.web-proxy.address	20888	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
NodeManager	yarn.nodemanager.address	8041	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
NodeManager	yarn.node manager.l ocalizer.address	8040	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
NodeManager	yarn.node manager.w ebapp.address	8044	TLS	emr-7.3.0+
NodeManager	mapreduce .shuffle.port[1]	13562	TLS	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
NodeManager	spark.shu ffle.service.port[2]	7337	Cifrado basado en AES de Spark	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+

[1] `mapreduce.shuffle.port` está alojado en YARN NodeManager pero solo lo usa MapReduce Hadoop.

[2] `spark.shuffle.service.port` está alojado en YARN NodeManager pero solo lo usa Apache Spark.

Problema conocido

Actualmente, la `yarn.log.server.url` configuración utiliza HTTP con el puerto 19888, lo que impide el acceso a los registros de las aplicaciones desde la interfaz de usuario del administrador de recursos. Afecta a la versión: EMR - 7.3.0 a EMR - 7.8.0.

Utilice la siguiente solución alternativa:

1. Modifique la `yarn.log.server.url` configuración `yarn-site.xml` para usar el `HTTPS` protocolo y el número `19890` de puerto.
2. Reinicie el administrador de recursos de YARN:`sudo systemctl restart hadoop-yarn-resourcemanager.service`.

HDFS de Hadoop

El nodo de nombre, el nodo de datos y el nodo de diario de Hadoop admiten TLS de forma predeterminada si el cifrado en tránsito está habilitado en los clústeres de EMR.

El Secure Hadoop RPC se ha establecido en `privacy` y usa el cifrado en tránsito basado en SASL. Esto requiere que la autenticación Kerberos esté habilitada en la configuración de seguridad.

Se recomienda no cambiar los puertos predeterminados que se utilizan para los puntos de conexión HTTPS.

El cifrado de datos en la transferencia de bloques de HDFS utiliza AES 256 y requiere que el cifrado en reposo esté habilitado en la configuración de seguridad.

Para obtener más información, consulte Hadoop in Secure Mode en la documentación de Apache Hadoop:

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Namenode	dfs.namenode.https-address	9871	TLS	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
Namenode	dfs.namenode.rpc-address	8020	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
Datanode		9865	TLS	

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
	dfs.datanode.https.address			emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
Datanode	dfs.datanode.address	9866	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
Nodo de diario	dfs.journalnode.https-address	8481	TLS	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+
Nodo de diario	dfs.journalnode.rpc-address	8485	SASL + Kerberos	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
DFSZKFailoverControlador	dfs.ha.zkfc.port	8019	Ninguno	El TLS para ZKFC solo se admite en Hadoop 3.4.0. Consulte HADOOP-18919 para más información. La versión 7.1.0 de Amazon EMR se encuentra actualmente en Hadoop 3.3.6. En el futuro, las versiones posteriores de Amazon EMR estarán en Hadoop 3.4.0

Hadoop MapReduce

Hadoop MapReduce, el servidor de historial de trabajos y MapReduce Shuffle admiten TLS de forma predeterminada cuando el cifrado en tránsito está habilitado en los clústeres de EMR.

La mezcla cifrada de Hadoop utiliza TLS. MapReduce

Se recomienda no cambiar los puertos predeterminados para los puntos de conexión HTTPS.

Para obtener más información, consulte Hadoop in Secure Mode en la documentación de Apache Hadoop:

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
JobHistoryServer	mapreduce .jobhisto ry.webapp .https.address	198-90	TLS	emr-7.3.0+
HILO NodeManager	mapreduce .shuffle.port[1]	13562	TLS	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+, emr-7.0.0+

[1] `mapreduce.shuffle.port` está alojado en YARN NodeManager pero solo lo usa MapReduce Hadoop.

Presto

En las versiones 5.6.0 y posteriores de Amazon EMR, la comunicación interna entre el coordinador de Presto y los trabajadores utiliza TLS. Amazon EMR establece todas las configuraciones necesarias para permitir una comunicación interna segura en Presto.

Si el conector utiliza el metaalmacén de Hive como almacén de metadatos, la comunicación entre el comunicador y el metaalmacén de Hive también se cifra con TLS.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Coordinador de Presto	http-server.https. port	8446	TLS	emr-5.6.0+, emr-6.0.0+, emr-7.0.0+
		8446	TLS	

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Trabajador de Presto	http-server.https.port			emr-5.6.0+, emr-6.0.0+, emr-7.0.0+

Trino

En las versiones 6.1.0 y posteriores de Amazon EMR, la comunicación interna entre el coordinador y los trabajadores de Presto utiliza TLS. Amazon EMR establece todas las configuraciones necesarias para permitir una comunicación interna segura en Trino.

Si el conector utiliza el metaalmacén de Hive como almacén de metadatos, la comunicación entre el comunicador y el metaalmacén de Hive también se cifra con TLS.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Coordinador de Trino	http-server.https.port	8446	TLS	emr-6.1.0+, emr-7.0.0+
Trabajador de Trino	http-server.https.port	8446	TLS	emr-6.1.0+, emr-7.0.0+

Hive y Tez

De forma predeterminada, el servidor Hive 2, el servidor de metaalmacén Hive, la IU web Hive LLAP Daemon y Hive LLAP intercambian toda la compatibilidad TLS cuando el cifrado en tránsito está habilitado en los clústeres de EMR. Para obtener más información acerca de las configuraciones de Hive, consulte Propiedades de la configuración.

La IU de Tez alojada en el servidor Tomcat también está habilitada para HTTPS cuando el cifrado en tránsito está habilitado en el clúster de EMR. Sin embargo, HTTPS está desactivado para el servicio

de IU web de Tez AM, por lo que los usuarios de AM no tienen acceso al archivo del almacén de claves del agente de escucha SSL que lo abre. También puede habilitar este comportamiento con las configuraciones booleanas `tez.am.tez-ui.webservice.enable.ssl` y `tez.am.tez-ui.webservice.enable.client.auth`.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
HiveServer2.	hive.server2.thrift.port	10000	TLS	emr-6.9.0+, emr-7.0.0+
HiveServer2.	hive.server2.thrift.http.port	10001	TLS	emr-6.9.0+, emr-7.0.0+
HiveServer2.	hive.server2.webui.port	10002	TLS	emr-7.3.0+
HiveMetastoreServer	hive.metastore.port	9083	TLS	emr-7.3.0+
LLAP Daemon	hive.llap.daemon.yarn.shuffle.port	15551	TLS	emr-7.3.0+
LLAP Daemon	hive.llap.daemon.web.port	15002	TLS	emr-7.3.0+
LLAP Daemon	hive.llap.daemon.output.service.port	15003	Ninguno	Hive no admite el cifrado en tránsito para este punto de conexión

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
LLAP Daemon	hive.llap .manageme nt.rpc.port	15004	Ninguno	Hive no admite el cifrado en tránsito para este punto de conexión
LLAP Daemon	hive.llap.plugin.r pc.port	Dinámico	Ninguno	Hive no admite el cifrado en tránsito para este punto de conexión
LLAP Daemon	hive.llap .daemon.rpc.port	Dinámico	Ninguno	Hive no admite el cifrado en tránsito para este punto de conexión
Web HCat	templeton.port	50111	TLS	emr-7.3.0+
Maestro de aplicación de Tez	tez.am.cl ient.am.port-range tez.am.cl ient.am.port-range	Dinámico	Ninguno	Tez no admite el cifrado en tránsito para este punto de conexión

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Maestro de aplicación de Tez	tez.am.tez-ui.webservice.port-range	Dinámico	Ninguno	Está deshabilitado de forma predeterminada. Se puede habilitar mediante las configuraciones de Tez en emr-7.3.0+
Tarea de Tez	N/D: no se puede configurar	Dinámico	Ninguno	Tez no admite el cifrado en tránsito para este punto de conexión
Tez UI	Se puede configurar mediante el servidor Tomcat en el que está alojada la IU de Tez	8080	TLS	emr-7.3.0+

Flink

Los puntos de conexión REST de Apache Flink y la comunicación interna entre los procesos de Flink admiten TLS de forma predeterminada cuando se habilita el cifrado en tránsito en los clústeres de EMR.

security.ssl.internal.enabled está establecido en true y usa cifrado en tránsito para la comunicación interna entre los procesos de Flink. Si no desea el cifrado en tránsito para la comunicación interna, deshabilite esa configuración. Se recomienda usar la configuración predeterminada para garantizar la máxima seguridad.

Amazon EMR establece security.ssl.rest.enabled en true y utiliza el cifrado en tránsito para los puntos de conexión REST. Además, Amazon EMR también se establece historyserver.web.ssl.enabled en verdadero para utilizar la comunicación TLS con el servidor de historial de Flink. Si no desea el cifrado en tránsito para los puntos REST, deshabilite esas configuraciones. Se recomienda usar la configuración predeterminada para garantizar la máxima seguridad.

Amazon EMR utiliza security.ssl.algorithms para especificar la lista de cifrados que utilizan el cifrado basado en AES. Anule esta configuración para usar los cifrados que desee.

Para obtener más información, consulte Configuración SSL en la documentación de Flink.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Servidor de historial de Flink	historyserver.web.port	8082	TLS	emr-7.3.0+
Servidor Rest del administrador de trabajos	rest.bind-port rest.port	Dinámico	TLS	emr-7.3.0+

HBase

Amazon EMR establece Secure Hadoop RPC en. privacy HMaster y RegionServer utilice el cifrado en tránsito basado en SASL. Esto requiere que la autenticación de Kerberos esté habilitada en la configuración de seguridad.

Amazon EMR establece hbase.ssl.enabled en verdadero y usa TLS para los puntos de conexión de la IU. Si no desea utilizar TLS para la IU de los puntos de conexión, deshabilite esta

configuración. Se recomienda usar la configuración predeterminada para garantizar la máxima seguridad.

Amazon EMR establece `hbase.rest.ssl.enabled` y `hbase.thrift.ssl.enabled` y usa TLS para los puntos de conexión de los servidores REST y Thirft, respectivamente. Si no desea utilizar TLS para estos puntos de conexión, deshabilite esta configuración. Se recomienda usar la configuración predeterminada para garantizar la máxima seguridad.

A partir de EMR 7.6.0, TLS es compatible con los puntos finales. HMaster RegionServer Amazon EMR también establece `hbase.server.netty.tls.enabled` y. `hbase.client.netty.tls.enabled` Si no desea utilizar TLS para estos puntos de enlace, desactive esta configuración. Le recomendamos que utilice la configuración predeterminada, que proporciona cifrado y, por lo tanto, mayor seguridad. Para obtener más información, consulte la sección Seguridad a nivel de transporte (TLS) en la comunicación HBase RPC en la Guía de HBase referencia de Apache.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
HMaster	HMaster	16 000	SASL + Kerberos TLS	SASL + Kerberos en emr-4.8.0+, emr-5.0.0+, emr-6.0.0+ y emr-7.0.0+ TLS en emr-7.6.0+
HMaster	HMaster INTERFAZ DE USUARIO	16010	TLS	emr-7.3.0+
RegionServer	RegionServer	16020	SASL + Kerberos	SASL + Kerberos en

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
			TLS	emr-4.8.0+, emr-5.0.0+, emr-6.0.0+ y emr-7.0.0+ TLS en emr-7.6.0+
RegionServer	RegionServer Información	16030	TLS	emr-7.3.0+
HBase Servidor REST	Servidor Rest	8070	TLS	emr-7.3.0+
HBase Servidor REST	IU REST	8085	TLS	emr-7.3.0+
Servidor Thrift de HBase	Servidor Thrift	9090	TLS	emr-7.3.0+
Servidor Thrift de HBase	IU del servidor Thrift	9095	TLS	emr-7.3.0+

Phoenix

Si habilitó el cifrado en tránsito en su clúster de EMR, Phoenix Query Server admite la propiedad TLS de `phoenix.queryserver.tls.enabled`, que está establecida de forma predeterminada en `true`.

Para obtener más información, consulte Configuraciones relacionadas con HTTPS en la documentación de Phoenix Query Server.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Servidor de consultas	phoenix.queryserver.http.port	8765	TLS	emr-7.3.0+

Oozie

OOZIE-3673 está disponible en Amazon EMR si ejecuta Oozie en Amazon EMR 7.3.0 y versiones posteriores. Si necesita configurar protocolos SSL o TLS personalizados al ejecutar una acción de correo electrónico, puede establecer la propiedad `oozie.email.smtp.ssl.protocols` en el archivo `oozie-site.xml`. De forma predeterminada, si ha activado el cifrado en tránsito, Amazon EMR utiliza el protocolo TLS v1.3.

OOZIE-3677 y OOZIE-3674 están también disponible en Amazon EMR si ejecuta Oozie en Amazon EMR 7.3.0 y versiones posteriores. Esto le permite especificar las propiedades `keyStoreType` y `trustStoreType` en `oozie-site.xml`. OOZIE-3674 añade el parámetro `--insecure` al cliente de Oozie para que pueda ignorar los errores de certificado.

Oozie exige la verificación del nombre de host mediante TLS, lo que significa que cualquier certificado que utilice para el cifrado en tránsito debe cumplir los requisitos de verificación del nombre de host. Si el certificado no cumple los criterios, el clúster podría quedarse atascado en la fase de `oozie share lib update` en la que Amazon EMR aprovisiona el clúster. Se recomienda actualizar los certificados para asegurarse de que cumplen con la verificación del nombre de host. Sin embargo, si no puede actualizar los certificados, puede deshabilitar el SSL para Oozie estableciendo la propiedad de `oozie.https.enabled` en `false` en la configuración del clúster.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
EmbeddedOozieServer	oozie.https.port	11443	TLS	emr-7.3.0+

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
EmbeddedOozieServer	oozie.email.smtp.port	25	TLS	emr-7.3.0+

Hue

De forma predeterminada, Hue admite TLS cuando el cifrado en tránsito está habilitado en los clústeres de Amazon EMR. Para obtener más información sobre las configuraciones de Hue, consulte Configurar Hue con HTTPS o SSL.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Hue	http_port	8888	TLS	emr-7.4.0+

Livy

De forma predeterminada, Livy admite TLS cuando el cifrado en tránsito está habilitado en los clústeres de Amazon EMR. Para obtener más información sobre las configuraciones de Livy, consulte Habilitar HTTPS con Apache Livy.

A partir de Amazon EMR 7.3.0, si usa el cifrado en tránsito y la autenticación Kerberos, no puede usar el servidor Livy para las aplicaciones Spark que dependen del metaalmacén de Hive. Este problema se solucionó en el HIVE-16340 y se resolvió por completo en el SPARK-44114 cuando la aplicación Spark de código abierto se pudo actualizar a Hive 3. Mientras tanto, puedes solucionar este problema si así lo deseas. `hive.metastore.use.SSL false` Para obtener más información, consulte Configuración de aplicaciones.

Para obtener más información, consulta cómo habilitar HTTPS con Apache Livy.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
livy-server	livy.server.port	8998	TLS	emr-7.4.0+

JupyterEnterpriseGateway

De forma predeterminada, Jupyter Enterprise Gateway admite TLS cuando el cifrado en tránsito está habilitado en los clústeres de Amazon EMR. Para obtener más información sobre las configuraciones de Jupyter Enterprise Gateway, consulte Proteger Enterprise Gateway Server.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
jupyter_enterprise_gateway	c. .port EnterpriseGatewayApp	9547	TLS	emr-7.4.0+

JupyterHub

De forma predeterminada, JupyterHub admite TLS cuando el cifrado en tránsito está habilitado en los clústeres de Amazon EMR. Para obtener más información, consulte Habilitar el cifrado SSL en la documentación. JupyterHub No se recomienda deshabilitar el cifrado.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
jupyter_hub	c. .port JupyterHub	9443	TLS	emr-5.14.0+, emr-6.0.0+, emr-7.0.0+

Zeppelin

De forma predeterminada, Zeppelin admite TLS cuando habilita el cifrado en tránsito en su clúster de EMR. Para obtener más información sobre las configuraciones de Zeppelin, consulte Configuración SSL en la documentación de Zeppelin.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Zeppelin	zeppelin. server.ssl.port	8890	TLS	7,30+

ZooKeeper

Amazon EMR se configura `serverCnxnFactory` para habilitar TLS `org.apache.zookeeper.server.NettyServerCnxnFactory` para el quórum de Zookeeper y la comunicación con el cliente.

`secureClientPort`especifica el puerto que escucha las conexiones TLS. Si el cliente no admite conexiones TLS a Zookeeper, los clientes pueden conectarse al puerto inseguro 2181 especificado en. `clientPort` Puede anular o deshabilitar estos dos puertos.

Amazon EMR establece ambos valores para habilitar la comunicación TLS `admin.forceHttps` `true` para el quórum `sslQuorum` y el servidor de administración. Si no desea el cifrado en tránsito para el quórum y el servidor de administración, puede deshabilitar esas configuraciones. Le recomendamos que utilice las configuraciones predeterminadas para garantizar la máxima seguridad.

Para obtener más información, consulte las opciones de cifrado, autenticación y autorización en la documentación de Zookeeper.

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
	secureClientPort	2281	TLS	emr-7.4.0+

Componente	punto de enlace	Puerto	Mecanismo de cifrado en tránsito	Compatible desde la versión
Servidor Zookeeper				
Servidor Zookeeper	Puertos de quórum	Hay 2: Los seguidores usan 2888 para conectarse con el líder. En la elección del líder se utiliza el 3888	TLS	emr-7.4.0+
Servidor Zookeeper	Admin.ServerPort	8341	TLS	emr-7.4.0+

AWS Identity and Access Management para Amazon EMR

AWS Identity and Access Management (IAM) es una herramienta Servicio de AWS que ayuda al administrador a controlar de forma segura el acceso a los AWS recursos. Los administradores de IAM controlan quién se puede autenticar (iniciar sesión) y autorizar (tener permisos) para utilizar los recursos de Amazon EMR. La IAM es una Servicio de AWS herramienta que puede utilizar sin coste adicional.

Temas

- Público
- Autenticación con identidades
- Administración de acceso mediante políticas
- Cómo funciona Amazon EMR con IAM
- Roles en tiempo de ejecución para los pasos de Amazon EMR

- Configuración de los roles de servicio de IAM de los permisos de Amazon EMR para los servicios y recursos de AWS
- Ejemplos de políticas de Amazon EMR basadas en identidades

Público

La forma de usar AWS Identity and Access Management (IAM) varía según el trabajo que realice en Amazon EMR.

Usuario de servicio: si utiliza el servicio Amazon EMR para realizar el trabajo, el administrador proporciona las credenciales y los permisos que necesita. A medida que utilice más características de Amazon EMR para realizar el trabajo, es posible que necesite permisos adicionales. Entender cómo se administra el acceso puede ayudarle a solicitar los permisos correctos al administrador. Si no puede acceder a una característica de Amazon EMR, consulte Solución de problemas de identidad y acceso de Amazon EMR.

Administrador de servicio: si está a cargo de los recursos de Amazon EMR de la empresa, probablemente tenga acceso completo a Amazon EMR. El trabajo consiste en determinar a qué características y recursos de Amazon EMR deben acceder los usuarios del servicio. Luego, debe enviar solicitudes a su gestionador de IAM para cambiar los permisos de los usuarios de su servicio. Revise la información de esta página para conocer los conceptos básicos de IAM. Para obtener más información acerca de cómo la empresa puede utilizar IAM con Amazon EMR, consulte Cómo funciona Amazon EMR con IAM.

Administrador de IAM: si es un administrador de IAM, es posible que desee obtener información sobre cómo escribir políticas para administrar el acceso a Amazon EMR. Para consultar ejemplos de políticas de Amazon EMR basadas en identidades que puede utilizar en IAM, consulte Ejemplos de políticas de Amazon EMR basadas en identidades.

Autenticación con identidades

La autenticación es la forma de iniciar sesión AWS con sus credenciales de identidad. Debe estar autenticado (con quien haya iniciado sesión AWS) como usuario de IAM o asumiendo una función de IAM. Usuario raíz de la cuenta de AWS

Puede iniciar sesión AWS como una identidad federada mediante las credenciales proporcionadas a través de una fuente de identidad. AWS IAM Identity Center Los usuarios (IAM Identity Center), la autenticación de inicio de sesión único de su empresa y sus credenciales de Google o Facebook son

ejemplos de identidades federadas. Al iniciar sesión como una identidad federada, su gestionador habrá configurado previamente la federación de identidades mediante roles de IAM. Cuando accedes AWS mediante la federación, asumes un rol de forma indirecta.

Según el tipo de usuario que sea, puede iniciar sesión en el portal AWS Management Console o en el de AWS acceso. Para obtener más información sobre cómo iniciar sesión AWS, consulte Cómo iniciar sesión Cuenta de AWS en su Guía del AWS Sign-In usuario.

Si accede AWS mediante programación, AWS proporciona un kit de desarrollo de software (SDK) y una interfaz de línea de comandos (CLI) para firmar criptográficamente sus solicitudes con sus credenciales. Si no utilizas AWS herramientas, debes firmar las solicitudes tú mismo. Para obtener más información sobre la firma de solicitudes, consulte AWS Signature Versión 4 para solicitudes API en la Guía del usuario de IAM.

Independientemente del método de autenticación que use, es posible que deba proporcionar información de seguridad adicional. Por ejemplo, le AWS recomienda que utilice la autenticación multifactor (MFA) para aumentar la seguridad de su cuenta. Para obtener más información, consulte Autenticación multifactor en la Guía del usuario de AWS IAM Identity Center y Autenticación multifactor AWS en IAM en la Guía del usuario de IAM.

Cuenta de AWS usuario root

Al crear una Cuenta de AWS, comienza con una identidad de inicio de sesión que tiene acceso completo a todos Servicios de AWS los recursos de la cuenta. Esta identidad se denomina usuario Cuenta de AWS raíz y se accede a ella iniciando sesión con la dirección de correo electrónico y la contraseña que utilizaste para crear la cuenta. Recomendamos encarecidamente que no utiliza el usuario raíz para sus tareas diarias. Proteja las credenciales del Usuario raíz y utilícelas solo para las tareas que solo el Usuario raíz pueda realizar. Para ver la lista completa de las tareas que requieren que inicie sesión como usuario raíz, consulta Tareas que requieren credenciales de usuario raíz en la Guía del usuario de IAM.

Identidad federada

Como práctica recomendada, exija a los usuarios humanos, incluidos los que requieren acceso de administrador, que utilicen la federación con un proveedor de identidades para acceder Servicios de AWS mediante credenciales temporales.

Una identidad federada es un usuario del directorio de usuarios de su empresa, un proveedor de identidades web AWS Directory Service, el directorio del Centro de Identidad o cualquier usuario al que acceda Servicios de AWS mediante las credenciales proporcionadas a través de una fuente de

identidad. Cuando las identidades federadas acceden Cuentas de AWS, asumen funciones y las funciones proporcionan credenciales temporales.

Para una administración de acceso centralizada, le recomendamos que utiliza AWS IAM Identity Center. Puede crear usuarios y grupos en el Centro de identidades de IAM, o puede conectarse y sincronizarse con un conjunto de usuarios y grupos de su propia fuente de identidad para usarlos en todas sus Cuentas de AWS aplicaciones. Para obtener más información, consulta ¿Qué es el Centro de identidades de IAM? en la Guía del usuario de AWS IAM Identity Center .

Usuarios y grupos de IAM

Un usuario de IAM es una identidad propia Cuenta de AWS que tiene permisos específicos para una sola persona o aplicación. Siempre que sea posible, recomendamos emplear credenciales temporales, en lugar de crear usuarios de IAM que tengan credenciales de larga duración como contraseñas y claves de acceso. No obstante, si tiene casos de uso específicos que requieran credenciales de larga duración con usuarios de IAM, recomendamos rotar las claves de acceso. Para más información, consulte Rotar las claves de acceso periódicamente para casos de uso que requieran credenciales de larga duración en la Guía del usuario de IAM.

Un grupo de IAM es una identidad que especifica un conjunto de usuarios de IAM. No puedes iniciar sesión como grupo. Puedes usar los grupos para especificar permisos para varios usuarios a la vez. Los grupos facilitan la administración de los permisos para grandes conjuntos de usuarios. Por ejemplo, puede asignar un nombre a un grupo IAMAdminsy concederle permisos para administrar los recursos de IAM.

Los usuarios son diferentes de los roles. Un usuario se asocia exclusivamente a una persona o aplicación, pero la intención es que cualquier usuario pueda asumir un rol que necesite. Los usuarios tienen credenciales de larga duración permanentes; no obstante, los roles proporcionan credenciales temporales. Para obtener más información, consulte Casos de uso para usuarios de IAM en la Guía del usuario de IAM.

Roles de IAM

Un rol de IAM es una identidad dentro de usted Cuenta de AWS que tiene permisos específicos. Es similar a un usuario de IAM, pero no está asociado a una persona determinada. Para asumir temporalmente un rol de IAM en el AWS Management Console, puede cambiar de un rol de usuario a uno de IAM (consola). Puedes asumir un rol llamando a una operación de AWS API AWS CLI o usando una URL personalizada. Para más información sobre los métodos para el uso de roles, consulta Métodos para asumir un rol en la Guía del usuario de IAM.

Los roles de IAM con credenciales temporales son útiles en las siguientes situaciones:

- Acceso de usuario federado: para asignar permisos a una identidad federada, puede crear un rol y definir sus permisos. Cuando se autentica una identidad federada, se asocia la identidad al rol y se le conceden los permisos define el rol. Para obtener información acerca de roles de federación, consulte Crear un rol para un proveedor de identidad de terceros (federación) en la Guía de usuario de IAM. Si utiliza el IAM Identity Center, debe configurar un conjunto de permisos. IAM Identity Center correlaciona el conjunto de permisos con un rol en IAM para controlar a qué pueden acceder las identidades después de autenticarse. Para obtener información acerca de los conjuntos de permisos, consulta Conjuntos de permisos en la Guía del usuario de AWS IAM Identity Center .
- Permisos de usuario de IAM temporales: un usuario de IAM puede asumir un rol de IAM para recibir temporalmente permisos distintos que le permitan realizar una tarea concreta.
- Acceso entre cuentas: puede utilizar un rol de IAM para permitir que alguien (una entidad principal de confianza) de otra cuenta acceda a los recursos de la cuenta. Los roles son la forma principal de conceder acceso entre cuentas. Sin embargo, con algunas Servicios de AWS, puedes adjuntar una política directamente a un recurso (en lugar de usar un rol como proxy). Para obtener información acerca de la diferencia entre los roles y las políticas basadas en recursos para el acceso entre cuentas, consulta Acceso a recursos entre cuentas en IAM en la Guía del usuario de IAM.
- Acceso entre servicios: algunos Servicios de AWS utilizan funciones en otros Servicios de AWS. Por ejemplo, cuando realizas una llamada en un servicio, es habitual que ese servicio ejecute aplicaciones en Amazon EC2 o almacene objetos en Amazon S3. Es posible que un servicio haga esto usando los permisos de la entidad principal, usando un rol de servicio o usando un rol vinculado al servicio.
  - Sesiones de acceso directo (FAS): cuando utilizas un usuario o un rol de IAM para realizar acciones en AWS ellas, se te considera principal. Cuando utiliza algunos servicios, es posible que realice una acción que desencadene otra acción en un servicio diferente. El FAS utiliza los permisos del principal que llama Servicio de AWS y los solicita Servicio de AWS para realizar solicitudes a los servicios descendentes. Las solicitudes de FAS solo se realizan cuando un servicio recibe una solicitud que requiere interacciones con otros Servicios de AWS recursos para completarse. En este caso, debe tener permisos para realizar ambas acciones. Para obtener información sobre las políticas a la hora de realizar solicitudes de FAS, consulte Reenviar sesiones de acceso.
  - Rol de servicio: un rol de servicio es un rol de IAM que adopta un servicio para realizar acciones en su nombre. Un administrador de IAM puede crear, modificar y eliminar un rol de servicio

desde IAM. Para obtener más información, consulte Creación de un rol para delegar permisos a un Servicio de AWS en la Guía del usuario de IAM.

- Función vinculada al servicio: una función vinculada a un servicio es un tipo de función de servicio que está vinculada a un. Servicio de AWS El servicio puede asumir el rol para realizar una acción en su nombre. Los roles vinculados al servicio aparecen en usted Cuenta de AWS y son propiedad del servicio. Un administrador de IAM puede ver, pero no editar, los permisos de los roles vinculados a servicios.

- Aplicaciones que se ejecutan en Amazon EC2: puedes usar un rol de IAM para administrar las credenciales temporales de las aplicaciones que se ejecutan en una EC2 instancia y realizan AWS CLI solicitudes a la AWS API. Esto es preferible a almacenar las claves de acceso en la EC2 instancia. Para asignar un AWS rol a una EC2 instancia y ponerlo a disposición de todas sus aplicaciones, debe crear un perfil de instancia adjunto a la instancia. Un perfil de instancia contiene el rol y permite que los programas que se ejecutan en la EC2 instancia obtengan credenciales temporales. Para obtener más información, consulte Usar un rol de IAM para conceder permisos a las aplicaciones que se ejecutan en EC2 instancias de Amazon en la Guía del usuario de IAM.

Administración de acceso mediante políticas

El acceso se controla AWS creando políticas y adjuntándolas a AWS identidades o recursos. Una política es un objeto AWS que, cuando se asocia a una identidad o un recurso, define sus permisos. AWS evalúa estas políticas cuando un director (usuario, usuario raíz o sesión de rol) realiza una solicitud. Los permisos en las políticas determinan si la solicitud se permite o se deniega. La mayoría de las políticas se almacenan AWS como documentos JSON. Para obtener más información sobre la estructura y el contenido de los documentos de política JSON, consulte Información general de políticas JSON en la Guía del usuario de IAM.

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

De forma predeterminada, los usuarios y los roles no tienen permisos. Un administrador de IAM puede crear políticas de IAM para conceder permisos a los usuarios para realizar acciones en los recursos que necesitan. A continuación, el administrador puede añadir las políticas de IAM a roles y los usuarios pueden asumirlos.

Las políticas de IAM definen permisos para una acción independientemente del método que se utiliza para realizar la operación. Por ejemplo, suponga que dispone de una política que permite la acción

`iam:GetRole`. Un usuario con esa política puede obtener información sobre el rol de la API AWS Management Console AWS CLI, la o la AWS API.

Políticas basadas en identidades

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede asociar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y en qué condiciones. Para obtener más información sobre cómo crear una política basada en identidad, consulte Creación de políticas de IAM en la Guía del usuario de IAM.

Las políticas basadas en identidades pueden clasificarse además como políticas insertadas o políticas administradas. Las políticas insertadas se integran directamente en un único usuario, grupo o rol. Las políticas administradas son políticas independientes que puede adjuntar a varios usuarios, grupos y roles de su Cuenta de AWS empresa. Las políticas administradas incluyen políticas AWS administradas y políticas administradas por el cliente. Para obtener más información sobre cómo elegir una política administrada o una política insertada, consulte Elegir entre políticas administradas y políticas insertadas en la Guía del usuario de IAM.

Políticas basadas en recursos

Las políticas basadas en recursos son documentos de política JSON que se asocian a un recurso. Los ejemplos de políticas basadas en recursos son las políticas de confianza de roles de IAM y las políticas de bucket de Amazon S3. En los servicios que admiten políticas basadas en recursos, los administradores de servicios pueden utilizarlos para controlar el acceso a un recurso específico. Para el recurso al que se asocia la política, la política define qué acciones puede realizar una entidad principal especificada en ese recurso y en qué condiciones. Debe especificar una entidad principal en una política en función de recursos. Los principales pueden incluir cuentas, usuarios, roles, usuarios federados o. Servicios de AWS

Las políticas basadas en recursos son políticas insertadas que se encuentran en ese servicio. No puedes usar políticas AWS gestionadas de IAM en una política basada en recursos.

Listas de control de acceso () ACLs

Las listas de control de acceso (ACLs) controlan qué responsables (miembros de la cuenta, usuarios o roles) tienen permisos para acceder a un recurso. ACLs son similares a las políticas basadas en recursos, aunque no utilizan el formato de documento de políticas JSON.

Amazon S3 y Amazon VPC son ejemplos de servicios compatibles. AWS WAF ACLs Para obtener más información ACLs, consulte la descripción general de la lista de control de acceso (ACL) en la Guía para desarrolladores de Amazon Simple Storage Service.

Otros tipos de políticas

AWS admite tipos de políticas adicionales y menos comunes. Estos tipos de políticas pueden establecer el máximo de permisos que los tipos de políticas más frecuentes le conceden.

- Límites de permisos: un límite de permisos es una característica avanzada que le permite establecer los permisos máximos que una política basada en identidad puede conceder a una entidad de IAM (usuario o rol de IAM). Puedes establecer un límite de permisos para una entidad. Los permisos resultantes son la intersección de las políticas basadas en la identidad de la entidad y los límites de permisos. Las políticas basadas en recursos que especifiquen el usuario o rol en el campo `Principal` no estarán restringidas por el límite de permisos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para obtener más información sobre los límites de los permisos, consulte Límites de permisos para las entidades de IAM en la Guía del usuario de IAM.
- Políticas de control de servicios (SCPs): SCPs son políticas de JSON que especifican los permisos máximos para una organización o unidad organizativa (OU). AWS Organizations AWS Organizations es un servicio para agrupar y administrar de forma centralizada varios de los Cuentas de AWS que son propiedad de su empresa. Si habilitas todas las funciones de una organización, puedes aplicar políticas de control de servicios (SCPs) a una o a todas tus cuentas. El SCP limita los permisos de las entidades en las cuentas de los miembros, incluidas las de cada una Usuario raíz de la cuenta de AWS. Para obtener más información sobre Organizations SCPs, consulte las políticas de control de servicios en la Guía del AWS Organizations usuario.
- Políticas de control de recursos (RCPs): RCPs son políticas de JSON que puedes usar para establecer los permisos máximos disponibles para los recursos de tus cuentas sin actualizar las políticas de IAM asociadas a cada recurso que poseas. El RCP limita los permisos de los recursos en las cuentas de los miembros y puede afectar a los permisos efectivos de las identidades, incluidos los permisos Usuario raíz de la cuenta de AWS, independientemente de si pertenecen a su organización. Para obtener más información sobre Organizations e RCPs incluir una lista de Servicios de AWS ese apoyo RCPs, consulte Políticas de control de recursos (RCPs) en la Guía del AWS Organizations usuario.
- Políticas de sesión: las políticas de sesión son políticas avanzadas que se pasan como parámetro cuando se crea una sesión temporal mediante programación para un rol o un usuario federado. Los permisos de la sesión resultantes son la intersección de las políticas basadas en identidades

del rol y las políticas de la sesión. Los permisos también pueden proceder de una política en función de recursos. Una denegación explícita en cualquiera de estas políticas anulará el permiso. Para más información, consulte Políticas de sesión en la Guía del usuario de IAM.

Varios tipos de políticas

Cuando se aplican varios tipos de políticas a una solicitud, los permisos resultantes son más complicados de entender. Para saber cómo se AWS determina si se debe permitir una solicitud cuando se trata de varios tipos de políticas, consulte la lógica de evaluación de políticas en la Guía del usuario de IAM.

Cómo funciona Amazon EMR con IAM

Antes de utilizar IAM para administrar el acceso a Amazon EMR, obtenga información sobre qué características de IAM se encuentran disponibles con Amazon EMR.

Características de IAM que puede utilizar con Amazon EMR

Característica de IAM	Compatibilidad con Amazon EMR
Políticas basadas en identidades	Sí
Políticas basadas en recursos	Sí
Acciones de políticas	Sí
Recursos de políticas	Sí
Claves de condición de política	Sí
ACLs	No
ABAC (etiquetas en políticas)	Sí
Credenciales temporales	Sí
Permisos de entidades principales	Sí
Roles de servicio	No
Roles vinculados al servicio	Sí

Para obtener una visión general de cómo funcionan Amazon EMR y otros AWS servicios con la mayoría de las funciones de IAM, consulte los AWS servicios que funcionan con IAM en la Guía del usuario de IAM.

Políticas de Amazon EMR basadas en identidades

Compatibilidad con las políticas basadas en identidad: sí

Las políticas basadas en identidad son documentos de políticas de permisos JSON que puede asociar a una identidad, como un usuario de IAM, un grupo de usuarios o un rol. Estas políticas controlan qué acciones pueden realizar los usuarios y los roles, en qué recursos y en qué condiciones. Para obtener más información sobre cómo crear una política basada en identidad, consulte Creación de políticas de IAM en la Guía del usuario de IAM.

Con las políticas basadas en identidades de IAM, puede especificar las acciones y los recursos permitidos o denegados, así como las condiciones en las que se permiten o deniegan las acciones. No es posible especificar la entidad principal en una política basada en identidad porque se aplica al usuario o rol al que está asociada. Para obtener más información sobre los elementos que puede utilizar en una política de JSON, consulte Referencia de los elementos de las políticas de JSON de IAM en la Guía del usuario de IAM.

Ejemplos de políticas basadas en identidades para Amazon EMR

Para ver ejemplos de políticas basadas en identidades de Amazon EMR, consulte Ejemplos de políticas de Amazon EMR basadas en identidades.

Políticas basadas en recursos de Amazon EMR

Compatibilidad con las políticas basadas en recursos: sí

Las políticas basadas en recursos son documentos de política JSON que se asocian a un recurso. Los ejemplos de políticas basadas en recursos son las políticas de confianza de roles de IAM y las políticas de bucket de Amazon S3. En los servicios que admiten políticas basadas en recursos, los administradores de servicios pueden utilizarlos para controlar el acceso a un recurso específico. Para el recurso al que se asocia la política, la política define qué acciones puede realizar una entidad principal especificada en ese recurso y en qué condiciones. Debe especificar una entidad principal en una política en función de recursos. Los principales pueden incluir cuentas, usuarios, roles, usuarios federados o. Servicios de AWS

Para habilitar el acceso entre cuentas, puede especificar toda una cuenta o entidades de IAM de otra cuenta como la entidad principal de una política en función de recursos. Añadir a una política en función de recursos una entidad principal entre cuentas es solo una parte del establecimiento de una relación de confianza. Cuando el principal y el recurso son diferentes Cuentas de AWS, el administrador de IAM de la cuenta de confianza también debe conceder a la entidad principal (usuario o rol) permiso para acceder al recurso. Para conceder el permiso, adjunte la entidad a una política basada en identidad. Sin embargo, si la política basada en recursos concede acceso a una entidad principal de la misma cuenta, no es necesaria una política basada en identidad adicional. Para obtener más información, consulte Cross account resource access in IAM en la Guía del usuario de IAM.

Acciones de políticas para Amazon EMR

Compatibilidad con las acciones de políticas: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Action` de una política JSON describe las acciones que puede utilizar para conceder o denegar el acceso en una política. Las acciones políticas suelen tener el mismo nombre que la operación de AWS API asociada. Hay algunas excepciones, como acciones de solo permiso que no tienen una operación de API coincidente. También hay algunas operaciones que requieren varias acciones en una política. Estas acciones adicionales se denominan acciones dependientes.

Incluya acciones en una política para conceder permisos y así llevar a cabo la operación asociada.

A fin de conocer una lista completa de acciones de políticas para Amazon EMR, consulte Acciones, recursos y claves de condición para Amazon EMR en la Referencia de autorizaciones de servicio.

Las acciones de políticas de Amazon EMR utilizan el siguiente prefijo antes de la acción:

```
EMR
```

Para especificar varias acciones en una única instrucción, sepárelas con comas.

```
"Action": [
      "EMR:action1",
      "EMR:action2"
         ]
```

Para ver ejemplos de políticas basadas en identidades de Amazon EMR, consulte Ejemplos de políticas de Amazon EMR basadas en identidades.

Recursos de políticas para Amazon EMR

Compatibilidad con los recursos de políticas: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Resource` de la política JSON especifica el objeto u objetos a los que se aplica la acción. Las instrucciones deben contener un elemento `Resource` o `NotResource`. Como práctica recomendada, especifique un recurso utilizando el Nombre de recurso de Amazon (ARN). Puedes hacerlo para acciones que admitan un tipo de recurso específico, conocido como permisos de nivel de recurso.

Para las acciones que no admiten permisos de nivel de recurso, como las operaciones de descripción, utiliza un carácter comodín (*) para indicar que la instrucción se aplica a todos los recursos.

```
"Resource": "*"
```

Para ver una lista de los tipos de recursos de Amazon EMR y sus tipos ARNs, consulte Recursos definidos por Amazon EMR en la Referencia de autorización de servicio. Para saber con qué acciones puede especificar el ARN de cada recurso, consulte Acciones, recursos y claves de condición para Amazon EMR.

Para ver ejemplos de políticas basadas en identidades de Amazon EMR, consulte Ejemplos de políticas de Amazon EMR basadas en identidades.

Claves de condición de políticas para Amazon EMR

Compatibilidad con claves de condición de políticas específicas del servicio: sí

Los administradores pueden usar las políticas de AWS JSON para especificar quién tiene acceso a qué. Es decir, qué entidad principal puede realizar acciones en qué recursos y en qué condiciones.

El elemento `Condition` (o bloque de `Condition`) permite especificar condiciones en las que entra en vigor una instrucción. El elemento `Condition` es opcional. Puedes crear expresiones

condicionales que utilizan operadores de condición, tales como igual o menor que, para que la condición de la política coincida con los valores de la solicitud.

Si especifica varios elementos de `Condition` en una instrucción o varias claves en un único elemento de `Condition`, AWS las evalúa mediante una operación `AND` lógica. Si especifica varios valores para una única clave de condición, AWS evalúa la condición mediante una `OR` operación lógica. Se deben cumplir todas las condiciones antes de que se concedan los permisos de la instrucción.

También puedes utilizar variables de marcador de posición al especificar condiciones. Por ejemplo, puedes conceder un permiso de usuario de IAM para acceder a un recurso solo si está etiquetado con su nombre de usuario de IAM. Para más información, consulta Elementos de la política de IAM: variables y etiquetas en la Guía del usuario de IAM.

AWS admite claves de condición globales y claves de condición específicas del servicio. Para ver todas las claves de condición AWS globales, consulte las claves de contexto de condición AWS globales en la Guía del usuario de IAM.

Para obtener una lista de las claves de condición de Amazon EMR y para más información sobre qué acciones y recursos admiten el uso de una clave de condición, consulte Acciones, recursos y claves de condición para Amazon EMR en la Referencia de autorizaciones de servicio.

Para ver ejemplos de políticas basadas en identidades de Amazon EMR, consulte Ejemplos de políticas de Amazon EMR basadas en identidades.

Listas de control de acceso (ACLs) en Amazon EMR

Soporta ACLs: No

Las listas de control de acceso (ACLs) controlan qué directores (miembros de la cuenta, usuarios o roles) tienen permisos para acceder a un recurso. ACLs son similares a las políticas basadas en recursos, aunque no utilizan el formato de documento de políticas JSON.

Control de acceso basado en atributos (ABAC) con Amazon EMR

Admite ABAC (etiquetas en las políticas)	Sí

El control de acceso basado en atributos (ABAC) es una estrategia de autorización que define permisos en función de atributos. En AWS, estos atributos se denominan etiquetas. Puede adjuntar

etiquetas a las entidades de IAM (usuarios o roles) y a muchos AWS recursos. El etiquetado de entidades y recursos es el primer paso de ABAC. A continuación, designa las políticas de ABAC para permitir operaciones cuando la etiqueta de la entidad principal coincida con la etiqueta del recurso al que se intenta acceder.

ABAC es útil en entornos que crecen con rapidez y ayuda en situaciones en las que la administración de las políticas resulta engorrosa.

Para controlar el acceso en función de etiquetas, debe proporcionar información de las etiquetas en el elemento de condición de una política utilizando las claves de condición `aws:ResourceTag/key-name`, `aws:RequestTag/key-name` o `aws:TagKeys`.

Si un servicio admite las tres claves de condición para cada tipo de recurso, el valor es Sí para el servicio. Si un servicio admite las tres claves de condición solo para algunos tipos de recursos, el valor es Parcial.

Para obtener más información sobre ABAC, consulte Definición de permisos con la autorización de ABAC en la Guía del usuario de IAM. Para ver un tutorial con los pasos para configurar ABAC, consulte Uso del control de acceso basado en atributos (ABAC) en la Guía del usuario de IAM.

Uso de credenciales temporales con Amazon EMR

Compatibilidad con credenciales temporales: sí

Algunas Servicios de AWS no funcionan cuando inicias sesión con credenciales temporales. Para obtener información adicional, incluidas las que Servicios de AWS funcionan con credenciales temporales, consulta Cómo Servicios de AWS funcionan con IAM en la Guía del usuario de IAM.

Utiliza credenciales temporales si inicia sesión en ellas AWS Management Console mediante cualquier método excepto un nombre de usuario y una contraseña. Por ejemplo, cuando accedes AWS mediante el enlace de inicio de sesión único (SSO) de tu empresa, ese proceso crea automáticamente credenciales temporales. También crea credenciales temporales de forma automática cuando inicia sesión en la consola como usuario y luego cambia de rol. Para obtener más información sobre el cambio de roles, consulte Cambio de un usuario a un rol de IAM (consola) en la Guía del usuario de IAM.

Puedes crear credenciales temporales manualmente mediante la AWS CLI API o. AWS A continuación, puede utilizar esas credenciales temporales para acceder AWS. AWS recomienda generar credenciales temporales de forma dinámica en lugar de utilizar claves de acceso a largo plazo. Para obtener más información, consulte Credenciales de seguridad temporales en IAM.

Permisos de entidades principales entre servicios de Amazon EMR

Admite sesiones de acceso directo (FAS): sí

Cuando utilizas un usuario o un rol de IAM para realizar acciones en AWSél, se te considera director. Cuando utiliza algunos servicios, es posible que realice una acción que desencadene otra acción en un servicio diferente. FAS utiliza los permisos del principal que llama y los que solicita Servicio de AWS para realizar solicitudes a los servicios descendentes. Servicio de AWS Las solicitudes de FAS solo se realizan cuando un servicio recibe una solicitud que requiere interacciones con otros Servicios de AWS recursos para completarse. En este caso, debe tener permisos para realizar ambas acciones. Para obtener información sobre las políticas a la hora de realizar solicitudes de FAS, consulte Reenviar sesiones de acceso.

Roles de servicio para Amazon EMR

Compatible con roles de servicio	No

Roles vinculado a servicios para Amazon EMR

Compatible con roles vinculados al servicio	Sí

Para más información sobre cómo crear o administrar roles vinculados a servicios, consulta Servicios de AWS que funcionan con IAM. Busque un servicio en la tabla que incluya `Yes` en la columna Rol vinculado a un servicio. Seleccione el vínculo Sí para ver la documentación acerca del rol vinculado a servicios para ese servicio.

Uso de etiquetas de clúster y cuaderno con políticas de IAM para el control de acceso

Es posible ajustar el permiso de las acciones de Amazon EMR asociadas con Cuadernos de Amazon EMR y clústeres de EMR mediante el control de acceso basado en etiquetas con políticas de IAM basadas en identidades. Puede utilizar claves de condición en un `Condition`elemento (también denominado bloque`Condition`) para permitir ciertas acciones solo cuando un bloc de notas, un clúster o ambos tengan una clave de etiqueta o una combinación clave-valor determinadas. También puede limitar la acción `CreateEditor` (que crea un cuaderno de EMR) y la acción `RunJobFlow` (que crea un clúster), de modo que una solicitud se envíe al crear dicho recurso.

En Amazon EMR, las claves de condición que se pueden utilizar en un elemento `Condition` se aplican únicamente a las acciones de la API de Amazon EMR donde `ClusterID` o `NotebookID` sean un parámetro de solicitud necesario. Por ejemplo, la ModifyInstanceGroupsacción no admite claves de contexto porque `ClusterID` es un parámetro opcional.

Al crear un cuaderno de EMR, se aplica una etiqueta predeterminada con una cadena de claves de `creatorUserId` que se convierte en el valor del ID de usuario de IAM que ha creado el cuaderno. Esto resulta útil para limitar las acciones permitidas en el bloc de notas únicamente al creador.

Las siguientes claves de condición están disponibles en Amazon EMR:

- Utilice la clave de contexto de condición `elasticmapreduce:ResourceTag/`*`TagKeyString`* para permitir o denegar las acciones de los usuarios en clústeres o blocs de notas con etiquetas que tienen la cadena *`TagKeyString`* que especifique. Si una acción pasa `ClusterID` y `NotebookID`, la condición se aplica al clúster y al bloc de notas. Esto significa que ambos recursos deben tener la cadena de clave de etiqueta o la combinación clave-valor que especifique. Puede utilizar el elemento `Resource` para limitar la instrucción de modo que se aplique únicamente a los clústeres o los blocs de notas que lo necesiten. Para obtener más información, consulte Ejemplos de políticas de Amazon EMR basadas en identidades.
- Use la clave de contexto de `elasticmapreduce:RequestTag/`*`TagKeyString`* condición para requerir una etiqueta específica con actions/API las llamadas. Por ejemplo, puede utilizar la clave de contexto de condición junto con la acción `CreateEditor` para requerir que se aplique una clave con la cadena *`TagKeyString`* a un bloc de notas cuando se crea este.

Ejemplos

Para ver una lista de las acciones de Amazon EMR, consulte Acciones definidas por Amazon EMR en la Guía del usuario de IAM.

Roles en tiempo de ejecución para los pasos de Amazon EMR

Un rol en tiempo de ejecución es un rol AWS Identity and Access Management (IAM) que puede especificar al enviar un trabajo o una consulta a un clúster de Amazon EMR. El trabajo o la consulta que envíe a su clúster de Amazon EMR utiliza el rol de tiempo de ejecución para acceder a AWS los recursos, como los objetos de Amazon S3. Puede especificar roles en tiempo de ejecución con Amazon EMR para los trabajos de Spark y Hive.

También puede especificar los roles en tiempo de ejecución al conectarse a los clústeres de Amazon EMR en Amazon SageMaker AI y al adjuntar un espacio de trabajo de Amazon EMR Studio a un clúster de EMR. Para obtener más información, consulte Conectarse a un clúster de Amazon EMR desde SageMaker AI Studio y. Ejecutar un espacio de trabajo de EMR Studio con un rol de tiempo de ejecución

Anteriormente, los clústeres de Amazon EMR ejecutaban trabajos o consultas de Amazon EMR con permisos basados en la política de IAM adjunta al perfil de instancia que utilizaba para lanzar el clúster. Esto significaba que las políticas tenían que contener la unión de todos los permisos para todos los trabajos y consultas que se ejecutaban en un clúster de Amazon EMR. Con los roles en tiempo de ejecución, ahora puede administrar el control de acceso de cada trabajo o consulta de forma individual, en lugar de compartir el perfil de instancia de Amazon EMR del clúster.

En los clústeres de Amazon EMR con funciones de tiempo de ejecución, también puede aplicar un control de acceso AWS Lake Formation basado a los trabajos y consultas de Spark, Hive y Presto en sus lagos de datos. Para obtener más información sobre cómo realizar la integración con AWS Lake Formation, consulte. Integre Amazon EMR con AWS Lake Formation

> **Note**
>
> Cuando especifica un rol de tiempo de ejecución para un paso de Amazon EMR, los trabajos o consultas que envíe solo pueden acceder a AWS los recursos que las políticas asociadas al rol de tiempo de ejecución permiten. Estos trabajos y consultas no pueden acceder al servicio de metadatos de instancias en las EC2 instancias del clúster ni utilizar el perfil de EC2 instancia del clúster para acceder a ningún AWS recurso.

Requisitos previos para lanzar un clúster de Amazon EMR con un rol en tiempo de ejecución

Temas

- Paso 1: configurar los controles de seguridad en Amazon EMR
- Paso 2: configurar un perfil de EC2 instancia para el clúster de Amazon EMR
- Paso 3: configurar una política de confianza

Paso 1: configurar los controles de seguridad en Amazon EMR

Usa la siguiente estructura JSON para crear una configuración de seguridad en AWS Command Line Interface (AWS CLI) y configúrala `EnableApplicationScopedIAMRole` en`true`. Para obtener más información acerca de las configuraciones de seguridad, consulte Uso de configuraciones de seguridad para configurar la seguridad del clúster de Amazon EMR.

```
{
    "AuthorizationConfiguration":{
        "IAMConfiguration":{
            "EnableApplicationScopedIAMRole":true
        }
    }
}
```

Le recomendamos que habilite siempre las opciones de cifrado en tránsito en la configuración de seguridad, de modo que los datos que se transfieran a través de Internet estén cifrados y no sean texto sin formato. Puede omitir estas opciones si no quiere conectarse a clústeres de Amazon EMR con funciones de tiempo de ejecución de SageMaker Runtime Studio o EMR Studio. Para configurar el cifrado de datos, consulte Configuración del cifrado de datos.

Como alternativa, puede crear una configuración de seguridad con ajustes personalizados con la AWS Management Console.

Paso 2: configurar un perfil de EC2 instancia para el clúster de Amazon EMR

Los clústeres de Amazon EMR utilizan el rol de perfil de EC2 instancia de Amazon para asumir los roles de tiempo de ejecución. Para usar roles en tiempo de ejecución con los pasos de Amazon EMR, agregue las siguientes políticas al rol de IAM que planea usar como rol de perfil de instancia. Para agregar políticas a un rol de IAM o editar una política integrada o administrada existente, consulte Adición y eliminación de permisos de identidad de IAM.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowRuntimeRoleUsage",
      "Effect": "Allow",
```

```
      "Action": [
        "sts:AssumeRole",
        "sts:TagSession"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/EMRRuntimeRole"
      ]
    }
  ]
}
```

Paso 3: configurar una política de confianza

Para cada rol de IAM que vaya a utilizar como rol en tiempo de ejecución, defina la siguiente política de confianza y sustituya `EMR_EC2_DefaultRole` por el rol de perfil de instancia. Para modificar la política de confianza de un rol de IAM, consulte Modificación de una política de confianza de rol.

```
{
    "Sid":"AllowAssumeRole",
    "Effect":"Allow",
    "Principal":{
        "AWS":"arn:aws:iam::<AWS_ACCOUNT_ID>:role/EMR_EC2_DefaultRole"
    },
    "Action":[
             "sts:AssumeRole",
             "sts:TagSession"
            ]
}
```

Lanzamiento de un clúster de Amazon EMR con un control de acceso basado en roles

Tras realizar las configuraciones, puede lanzar un clúster de Amazon EMR con la configuración de seguridad de Paso 1: configurar los controles de seguridad en Amazon EMR. Para usar las funciones de tiempo de ejecución con los pasos de Amazon EMR, utilice la etiqueta de lanzamiento `emr-6.7.0` o una versión posterior y seleccione Hive, Spark o ambas como aplicación de clúster. CloudWatchAgent es compatible con los clústeres de roles en tiempo de ejecución para EMR 7.6 y versiones posteriores. Para conectarte desde SageMaker AI Studio, usa la versión `emr-6.9.0` o una versión posterior y selecciona Livy, Spark, Hive o Presto como aplicación de clúster. Para obtener instrucciones sobre cómo lanzar el clúster, consulte Especificar una configuración de seguridad para un clúster de Amazon EMR.

Envío de trabajos de Spark siguiendo los pasos de Amazon EMR

A continuación, se muestra un ejemplo de cómo ejecutar el HdfsTest ejemplo incluido con Apache Spark. Esta llamada a la API solo se realiza correctamente si el rol en tiempo de ejecución de Amazon EMR proporcionado puede acceder a `S3_LOCATION`.

```
RUNTIME_ROLE_ARN=<runtime-role-arn>
S3_LOCATION=<s3-path>
REGION=<aws-region>
CLUSTER_ID=<cluster-id>

aws emr add-steps --cluster-id $CLUSTER_ID \
--steps '[{ "Name": "Spark Example", "ActionOnFailure": "CONTINUE", "Jar":"command-
runner.jar","Args" : ["spark-example","HdfsTest", "$S3_LOCATION"] }]' \
--execution-role-arn $RUNTIME_ROLE_ARN \
--region $REGION
```

> **Note**
>
> Le recomendamos que desactive el acceso SSH al clúster de Amazon EMR y que solo permita que la API de Amazon EMR `AddJobFlowSteps` acceda al clúster.

Envío de trabajos de Hive siguiendo los pasos de Amazon EMR

En el siguiente ejemplo, se utilizan los pasos de Apache Hive con Amazon EMR para enviar un trabajo y ejecutar el archivo `QUERY_FILE.hql`. Esta consulta solo se realiza correctamente si el rol en tiempo de ejecución proporcionado puede acceder a la ruta de Amazon S3 del archivo de consulta.

```
RUNTIME_ROLE_ARN=<runtime-role-arn>
REGION=<aws-region>
CLUSTER_ID=<cluster-id>

aws emr add-steps --cluster-id $CLUSTER_ID \
--steps '[{ "Name": "Run hive query using command-runner.jar - simple
 select","ActionOnFailure":"CONTINUE", "Jar": "command-runner.jar","Args" :["hive -
f","s3://DOC_EXAMPLE_BUCKET/QUERY_FILE.hql"] }]' \
--execution-role-arn $RUNTIME_ROLE_ARN \
--region $REGION
```

Conéctese a clústeres de Amazon EMR con funciones de tiempo de ejecución desde un bloc de notas de SageMaker AI Studio

Puede aplicar las funciones de tiempo de ejecución de Amazon EMR a las consultas que ejecute en los clústeres de Amazon EMR desde AI Studio. SageMaker Para hacerlo, siga estos pasos:

1. Siga las instrucciones de Launch Amazon SageMaker AI Studio para crear un estudio de SageMaker IA.
2. En la interfaz de usuario de SageMaker AI Studio, inicie un bloc de notas con núcleos compatibles. Por ejemplo, inicia una SparkMagic imagen con un PySpark núcleo.
3. Elija un clúster de Amazon EMR en SageMaker AI Studio y, a continuación, elija Connect.
4. Seleccione un rol en tiempo de ejecución y, a continuación, elija Conectar.

Esto creará una celda de bloc de notas de SageMaker IA con comandos mágicos para conectarse a su clúster de Amazon EMR con la función de tiempo de ejecución de Amazon EMR elegida. En la celda del cuaderno, puede introducir y ejecutar consultas con el rol en tiempo de ejecución y el control de acceso basado en Lake Formation. Para ver un ejemplo más detallado, consulte Aplicar controles de acceso a datos detallados con AWS Lake Formation Amazon EMR de Amazon AI Studio. SageMaker

Control del acceso al rol en tiempo de ejecución de Amazon EMR

Puede controlar el acceso al rol en tiempo de ejecución con la clave de condición `elasticmapreduce:ExecutionRoleArn`. La siguiente política permite a una entidad principal de IAM utilizar un rol de IAM denominado `Caller`, o cualquier rol de IAM que comience por la cadena `CallerTeamRole`, como rol en tiempo de ejecución.

> **Important**
>
> Debe crear una condición basada en la clave de `elasticmapreduce:ExecutionRoleArn` contexto cuando conceda a la persona que llama acceso para llamar al `AddJobFlowSteps` o `GetClusterSessionCredentials` APIs, como se muestra en el siguiente ejemplo.

```
{
    "Sid":"AddStepsWithSpecificExecRoleArn",
    "Effect":"Allow",
    "Action":[
```

```
        "elasticmapreduce:AddJobFlowSteps"
    ],
    "Resource":"*",
    "Condition":{
        "StringEquals":{
            "elasticmapreduce:ExecutionRoleArn":[
                "arn:aws:iam::<AWS_ACCOUNT_ID>:role/Caller"
            ]
        },
        "StringLike":{
            "elasticmapreduce:ExecutionRoleArn":[
                "arn:aws:iam::<AWS_ACCOUNT_ID>:role/CallerTeamRole*"
            ]
        }
    }
}
```

Establecimiento de confianza entre los roles en tiempo de ejecución y los clústeres de Amazon EMR

Amazon EMR genera un identificador único `ExternalId` para cada configuración de seguridad con la autorización del rol en tiempo de ejecución activada. Esta autorización permite a cada usuario ser propietario de un conjunto de roles en tiempo de ejecución para utilizarlos en los clústeres que les pertenecen. Por ejemplo, en una empresa, cada departamento puede usar su identificador externo para actualizar la política de confianza en su propio conjunto de roles en tiempo de ejecución.

Puede obtener el identificador externo con la API `DescribeSecurityConfiguration` de Amazon EMR de Amazon, tal y como se muestra en el ejemplo siguiente.

```
aws emr describe-security-configuration --name 'iamconfig-with-lf'{"Name": "iamconfig-
with-lf",
    "SecurityConfiguration":
        "{\"AuthorizationConfiguration\":{\"IAMConfiguration\":
{\"EnableApplicationScopedIAMRole\
        ":true,\"ApplicationScopedIAMRoleConfiguration\":{\"PropagateSourceIdentity
\":true,\"Exter
        nalId\":\"FXH5TSACFDWUCDSR3YQE2O7ETPUSM4OBCGLYWODSCUZDNZ4Y\"}},\"Lake
        FormationConfiguration\":{\"AuthorizedSessionTagValue\":\"Amazon EMR\"}}}",
    "CreationDateTime": "2022-06-03T12:52:35.308000-07:00"
}
```

Para obtener información sobre cómo usar un identificador externo, consulta Cómo usar un identificador externo al conceder acceso a tus AWS recursos a un tercero.

Auditoría

Para supervisar y controlar las acciones que realizan los usuarios finales con los roles de IAM, puede activar la característica de identidad de origen. Para obtener más información sobre la identidad de origen, consulte Monitoreo y control de las acciones realizadas con roles asumidos.

Para hacer un seguimiento de la identidad de origen, establezca `ApplicationScopedIAMRoleConfiguration/PropagateSourceIdentity` en `true` en la configuración de seguridad de la siguiente manera.

```
{
    "AuthorizationConfiguration":{
        "IAMConfiguration":{
            "EnableApplicationScopedIAMRole":true,
            "ApplicationScopedIAMRoleConfiguration":{
                "PropagateSourceIdentity":true
            }
        }
    }
}
```

Cuando establece `PropagateSourceIdentity` en `true`, Amazon EMR aplica la identidad de origen de las credenciales de llamada a un trabajo o sesión de consulta que cree con el rol en tiempo de ejecución. Si no hay identidad de origen en las credenciales de llamada, Amazon EMR no establece la identidad de origen.

Para usar esta propiedad, proporcione permisos `sts:SetSourceIdentity` a su perfil de instancia de la siguiente manera.

```
{ // PropagateSourceIdentity statement
    "Sid":"PropagateSourceIdentity",
    "Effect":"Allow",
    "Action":"sts:SetSourceIdentity",
    "Resource":[
        <runtime-role-ARN>
    ],
    "Condition":{
        "StringEquals":{
            "sts:SourceIdentity":<source-identity>
        }
    }
```

```
}
```

También debe agregar la instrucción `AllowSetSourceIdentity` a la política de confianza de sus roles en tiempo de ejecución.

```
{ // AllowSetSourceIdentity statement
    "Sid":"AllowSetSourceIdentity",
    "Effect":"Allow",
    "Principal":{
        "AWS":"arn:aws:iam::<AWS_ACCOUNT_ID>:role/EMR_EC2_DefaultRole"
    },
    "Action":[
        "sts:SetSourceIdentity",
        "sts:AssumeRole"
    ],
    "Condition":{
        "StringEquals":{
            "sts:SourceIdentity":<source-identity>
        }
    }
}
```

Consideraciones adicionales

> **Note**
>
> Con la versión Amazon EMR`emr-6.9.0`, es posible que experimente errores intermitentes al conectarse a los clústeres de Amazon EMR desde AI Studio. SageMaker Si desea solucionar este problema, puede instalar el parche con una acción de arranque al lanzar el clúster. Para obtener más información sobre el parche, consulte Problemas conocidos de la versión 6.9.0 de Amazon EMR.

Además, tenga en cuenta lo siguiente al configurar los roles en tiempo de ejecución para Amazon EMR.

- Amazon EMR admite roles en tiempo de ejecución en todas las Regiones de AWS comerciales.
- Los pasos de Amazon EMR admiten los trabajos de Apache Spark y Apache Hive con roles en tiempo de ejecución cuando se utiliza la versión `emr-6.7.0` o versiones posteriores.

- SageMaker AI Studio admite consultas de Spark, Hive y Presto con funciones de tiempo de ejecución cuando utilizas una versión o una versión posterior. `emr-6.9.0`
- Los siguientes núcleos de cuadernos de SageMaker AI admiten funciones de tiempo de ejecución:
  - DataScience — Núcleo de Python 3
  - DataScience 2.0 — Núcleo de Python 3
  - DataScience 3.0 — Núcleo de Python 3
  - SparkAnalytics 1.0 — SparkMagic y PySpark núcleos
  - SparkAnalytics 2.0 — SparkMagic y núcleos PySpark
  - SparkMagic — núcleo PySpark
- Amazon EMR admite pasos que utilizan `RunJobFlow` únicamente en el momento de la creación del clúster. Esta API no admite roles en tiempo de ejecución.
- Amazon EMR no admite roles en tiempo de ejecución en clústeres que usted configure para que tengan una alta disponibilidad.
- A partir de la versión 7.5.0 y posteriores de Amazon EMR, los roles de tiempo de ejecución permiten ver las interfaces de usuario de Spark y YARN (UIs), como las siguientes: Spark Live UI, Spark History Server NodeManager, YARN y YARN. ResourceManager Cuando navegues hasta ellas UIs, aparecerá un mensaje de nombre de usuario y contraseña. Los nombres de usuario y las contraseñas se pueden generar mediante el uso de la API EMR. GetClusterSessionCredentials Para obtener más información sobre los detalles de uso de la API, consulte. GetClusterSessionCredentials

  Un ejemplo de cómo utilizar la GetClusterSessionCredentials API EMR es el siguiente:

  ```
  aws emr  get-cluster-session-credentials --cluster-id <cluster_ID> --execution-role-
  arn <IAM_role_arn>
  ```

- Debe evitar los argumentos de los comandos de Bash cuando ejecute comandos con el archivo JAR de `command-runner.jar`:

  ```
  aws emr add-steps --cluster-id <cluster-id> --steps '[{"Name":"sample-
  step","ActionOnFailure":"CONTINUE","Jar":"command-runner.jar","Properties":"","Args":
  ["bash","-c","\"aws s3 ls\""],"Type":"CUSTOM_JAR"}]' --execution-role-
  arn <IAM_ROLE_ARN>
  ```

Además, debe evitar los argumentos de los comandos de Bash al ejecutar comandos con el ejecutor de scripts. El siguiente es un ejemplo que muestra cómo configurar las propiedades de Spark, con los caracteres de escape incluidos:

```
"\"--conf spark.sql.autoBroadcastJoinThreshold=-1\n--conf
 spark.cradle.RSv2Mode.enabled=true\""
```

- Los roles en tiempo de ejecución no permiten controlar el acceso a los recursos del clúster, como HDFS y HMS.
- Los roles de tiempo de ejecución no admiten docker/containers.

Configuración de los roles de servicio de IAM de los permisos de Amazon EMR para los servicios y recursos de AWS

Amazon EMR y las aplicaciones como Hadoop necesitan permisos para obtener acceso a otros recursos de AWS y realizar acciones cuando se ejecutan. Cada clúster de Amazon EMR debe tener un rol de servicio y un rol para el perfil de EC2 instancia de Amazon. Para obtener más información, consulte Roles de IAM y Uso de perfiles de instancia en la Guía del usuario de IAM. Las políticas de IAM asociadas a estos roles proporcionan permisos que permiten al clúster interoperar con otros servicios de AWS en nombre de un usuario.

Es necesario otro rol adicional, el rol de escalado automático, si el clúster utiliza el escalado automático en Amazon EMR. El rol AWS de servicio de EMR Notebooks es obligatorio si utiliza EMR Notebooks.

Amazon EMR proporciona roles predeterminados y políticas administradas predeterminadas que determinan permisos los permisos de cada rol. Las políticas administradas las crea y mantiene AWS, por lo que se actualizan automáticamente si cambian los requisitos del servicio. Para obtener más información, consulte Políticas administradas por AWS en la Guía del usuario de IAM.

Si está creando un clúster o un cuaderno por primera vez en una cuenta, los roles de Cuadernos de Amazon EMR aún no existen. Tras crearlas, puede ver las funciones, las políticas asociadas a ellas y los permisos permitidos o denegados por las políticas en la consola de IAM (https://console.aws.amazon.com/iam/). Puede especificar los roles predeterminados para Amazon EMR para crear y utilizar, puede crear sus propios roles y especificarlos de manera individual al crear un clúster para personalizar los permisos, y puede especificar roles predeterminados que se utilizarán

al crear un clúster mediante la AWS CLI. Para obtener más información, consulte Personalización de roles de IAM con Amazon EMR.

Modificación de políticas basadas en identidades para obtener permisos y así transferir roles de servicio para Amazon EMR

Las políticas administradas de forma predeterminada con todos los permisos de Amazon EMR incorporan configuraciones de seguridad `iam:PassRole`, entre las que se incluyen las siguientes:

- Permisos `iam:PassRole` solo para roles específicos de Amazon EMR predeterminados.
- `iam:PassedToService`condiciones que le permiten usar la política solo con AWS servicios específicos, como `elasticmapreduce.amazonaws.com` y`ec2.amazonaws.com`.

Puede ver la versión JSON de las políticas Amazon EMRFull AccessPolicy _v2 y Amazon EMRService Policy_v2 en la consola de IAM. Le recomendamos crear nuevos clústeres con las políticas administradas de la versión 2.

Resumen de rol de servicio

En la siguiente tabla se muestran los roles de servicio de IAM asociados con Amazon EMR para una consulta rápida.

Función	Rol predeterminado	Descripción	Política administrada predeterminada
Rol de servicio para Amazon EMR (rol de EMR)	`EMR_DefaultRole_V2`	Permite a Amazon EMR llamar a otros AWS servicios en su nombre al aprovisionar recursos y realizar acciones a nivel de servicio. Este rol es necesario para todos los clústeres.	`AmazonEMRServicePolicy_v2` **Important** La solicitud de instancias de spot requiere un rol vinculado a un servicio. Si este rol no

Función	Rol predeterminado	Descripción	Política administrada predeterminada
			existe, el rol de servicio de Amazon EMR debe tener permisos para crearlo; en caso contrario, se producirá un error de permisos. Si planea solicitar instancias de spot, debe actualizar esta política para incluir una instrucción que permita la creación de este rol vinculado a un servicio. Para obtener más información, consulte Rol de servicio para Amazon EMR (rol de EMR) un rol vinculado a un servicio para las solicitudes de instancias

Función	Rol predeterminado	Descripción	Política administrada predeterminada
			puntuales en la Guía EC2 del usuario de Amazon.

Función	Rol predeterminado	Descripción	Política administrada predeterminada
Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia)	`EMR_EC2_DefaultRole`	Los procesos de aplicación que se ejecutan sobre el ecosistema de Hadoop en instancias de clúster utilizan esta función cuando llaman a otros servicios. AWS Para obtener acceso a los datos en Amazon S3 usando EMRFS, puede especificar diferentes roles que se asumirán en función de la ubicación de los datos en Amazon S3. Por ejemplo, varios equipos pueden acceder a una única “cuenta de almacenamiento” de datos de Amazon S3. Para obtener más información, consulte Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3. Este rol es necesario para todos los clústeres.	`AmazonElasticMapReduceforEC2Role` . Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia).

Función	Rol predeterminado	Descripción	Política administrada predeterminada
Rol de servicio para el escalado automático en Amazon EMR (rol de Auto Scaling)	`EMR_AutoScaling_DefaultRole`	Permite acciones adicionales para entornos de escalado dinámico. Solo es obligatorio para los clústeres que utilizan el escalado automático en Amazon EMR. Para obtener más información, consulte Uso del escalado automático con una política personalizada para grupos de instancias en Amazon EMR.	`AmazonElasticMapReduceforAutoScalingRole` . Para obtener más información, consulte Rol de servicio para el escalado automático en Amazon EMR (rol de Auto Scaling).

Función	Rol predeterminado	Descripción	Política administrada predeterminada
Rol de servicio para Cuadernos de Amazon EMR	`EMR_Notebooks_DefaultRole`	Proporciona los permisos que un bloc de notas EMR necesita para acceder a otros AWS recursos y realizar acciones. Necesario solo si se utiliza Cuadernos de Amazon EMR.	`AmazonElasticMapReduceEditorsRole` . Para obtener más información, consulte Rol de servicio para Cuadernos de Amazon EMR. También se asocia `S3FullAccessPolicy` de forma predeterminada. El contenido de esta política se muestra a continuación. JSON `{ "Version": "2012-10-17", "Statement": [ { "Effect": "Allow", "Action": [`

Función	Rol predeterminado	Descripción	Política administrada predeterminada
			`"s3:*" ], "Resource": [ "*" ], "Sid": "AllowS3" } ] }`

Función	Rol predeterminado	Descripción	Política administrada predeterminada
Rol vinculado a servicio	`AWSServiceRoleForEMRCleanup`	Amazon EMR crea automáticamente un rol vinculado a un servicio. Si el servicio de Amazon EMR ha perdido la capacidad de limpiar los EC2 recursos de Amazon, Amazon EMR puede utilizar esta función para limpiarlos. Si un clúster utiliza instancias de spot, la política de permisos vinculada al Rol de servicio para Amazon EMR (rol de EMR) debe permitir la creación de un rol vinculado al servicio. Para obtener más información, consulte Uso de roles vinculados a servicios para Amazon EMR.	`AmazonEMRCleanupPolicy`

Temas

- Roles de servicio de IAM utilizados por Amazon EMR
- Personalización de roles de IAM con Amazon EMR
- Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3
- Uso de políticas basadas en recursos para el acceso de Amazon EMR a Catálogo de datos de AWS Glue

- Uso de roles de IAM con las aplicaciones que llaman directamente a los servicios de AWS
- Cómo permitir a los usuarios y grupos crear y modificar roles

Roles de servicio de IAM utilizados por Amazon EMR

Amazon EMR utiliza roles de servicio de IAM para llevar a cabo acciones en su nombre al aprovisionar los recursos del clúster, ejecutar aplicaciones, escalar recursos de forma dinámica y crear y ejecutar Cuadernos de Amazon EMR. Amazon EMR utiliza los siguientes roles cuando interactúa con otros servicios de AWS . Cada rol tiene una función exclusiva en Amazon EMR. Los temas de esta sección describen la función del rol y facilitan los roles y la política de permisos predeterminados para cada rol.

Si tiene un código de aplicación en su clúster que llama directamente a AWS los servicios, es posible que necesite usar el SDK para especificar las funciones. Para obtener más información, consulte Uso de roles de IAM con las aplicaciones que llaman directamente a los servicios de AWS.

Temas

- Rol de servicio para Amazon EMR (rol de EMR)
- Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia)
- Rol de servicio para el escalado automático en Amazon EMR (rol de Auto Scaling)
- Rol de servicio para Cuadernos de Amazon EMR
- Uso de roles vinculados a servicios para Amazon EMR

Rol de servicio para Amazon EMR (rol de EMR)

La función Amazon EMR define las acciones permitidas para Amazon EMR cuando aprovisiona recursos y realiza tareas de nivel de servicio que no se realizan en el contexto de una instancia de Amazon que se ejecuta dentro de un clúster. EC2 Por ejemplo, la función de servicio se utiliza para aprovisionar EC2 instancias cuando se lanza un clúster.

- El nombre del rol predeterminado es `EMR_DefaultRole_V2`.
- La política administrada predeterminada con ámbito de aplicación de Amazon EMR y asociada a `EMR_DefaultRole_V2` es `AmazonEMRServicePolicy_v2`. Esta política, versión 2, sustituye a la política administrada predeterminada, `AmazonElasticMapReduceRole`, ya obsoleta.

`AmazonEMRServicePolicy_v2` depende del acceso limitado a los recursos que Amazon EMR aprovisiona o utiliza. Cuando utilice esta política, tendrá que pasar la etiqueta de usuario `for-use-with-amazon-emr-managed-policies = true` al aprovisionar el clúster. Amazon EMR propagará automáticamente esas etiquetas. Además, es posible que tenga que añadir manualmente una etiqueta de usuario a tipos específicos de recursos, como grupos de EC2 seguridad que Amazon EMR no creó. Consulte Etiquetado de recursos para usar políticas administradas.

> **Important**
>
> Amazon EMR utiliza esta función de servicio de Amazon EMR y la `AWSServiceRoleForEMRCleanup` función para limpiar los recursos del clúster de su cuenta que ya no utiliza, como las instancias de Amazon. EC2 Debe incluir acciones para que las políticas de rol eliminen o terminen los recursos. De lo contrario, Amazon EMR no podrá realizar estas acciones de limpieza y podría incurrir en costos por los recursos no utilizados que permanecen en el clúster.

A continuación se muestra el contenido de la política `AmazonEMRServicePolicy_v2`. También puede ver el contenido actual de la política `AmazonEMRServicePolicy_v2`administrada en la consola de IAM.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "CreateInTaggedNetwork",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateNetworkInterface",
        "ec2:RunInstances",
        "ec2:CreateFleet",
        "ec2:CreateLaunchTemplate",
        "ec2:CreateLaunchTemplateVersion"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:subnet/*",
        "arn:aws:ec2:*:*:security-group/*"
      ],
```

```
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "CreateWithEMRTaggedLaunchTemplate",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateFleet",
          "ec2:RunInstances",
          "ec2:CreateLaunchTemplateVersion"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:launch-template/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "CreateEMRTaggedLaunchTemplate",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateLaunchTemplate"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:launch-template/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "CreateEMRTaggedInstancesAndVolumes",
        "Effect": "Allow",
        "Action": [
          "ec2:RunInstances",
          "ec2:CreateFleet"
```

```
        ],
        "Resource": [
          "arn:aws:ec2:*:*:instance/*",
          "arn:aws:ec2:*:*:volume/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "ResourcesToLaunchEC2",
        "Effect": "Allow",
        "Action": [
          "ec2:RunInstances",
          "ec2:CreateFleet",
          "ec2:CreateLaunchTemplate",
          "ec2:CreateLaunchTemplateVersion"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:network-interface/*",
          "arn:aws:ec2:*::image/ami-*",
          "arn:aws:ec2:*:*:key-pair/*",
          "arn:aws:ec2:*:*:capacity-reservation/*",
          "arn:aws:ec2:*:*:placement-group/pg-*",
          "arn:aws:ec2:*:*:fleet/*",
          "arn:aws:ec2:*:*:dedicated-host/*",
          "arn:aws:resource-groups:*:*:group/*"
        ]
      },
      {
        "Sid": "ManageEMRTaggedResources",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateLaunchTemplateVersion",
          "ec2:DeleteLaunchTemplate",
          "ec2:DeleteNetworkInterface",
          "ec2:ModifyInstanceAttribute",
          "ec2:TerminateInstances"
        ],
        "Resource": [
          "*"
        ],
```

```
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "ManageTagsOnEMRTaggedResources",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateTags",
          "ec2:DeleteTags"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:instance/*",
          "arn:aws:ec2:*:*:volume/*",
          "arn:aws:ec2:*:*:network-interface/*",
          "arn:aws:ec2:*:*:launch-template/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "CreateNetworkInterfaceNeededForPrivateSubnet",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateNetworkInterface"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:network-interface/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "TagOnCreateTaggedEMRResources",
        "Effect": "Allow",
        "Action": [
```

```
        "ec2:CreateTags"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:network-interface/*",
        "arn:aws:ec2:*:*:instance/*",
        "arn:aws:ec2:*:*:volume/*",
        "arn:aws:ec2:*:*:launch-template/*"
      ],
      "Condition": {
        "StringEquals": {
          "ec2:CreateAction": [
            "RunInstances",
            "CreateFleet",
            "CreateLaunchTemplate",
            "CreateNetworkInterface"
          ]
        }
      }
    },
    {
      "Sid": "TagPlacementGroups",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateTags",
        "ec2:DeleteTags"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:placement-group/pg-*"
      ]
    },
    {
      "Sid": "ListActionsForEC2Resources",
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeAccountAttributes",
        "ec2:DescribeCapacityReservations",
        "ec2:DescribeDhcpOptions",
        "ec2:DescribeImages",
        "ec2:DescribeInstances",
        "ec2:DescribeInstanceTypeOfferings",
        "ec2:DescribeLaunchTemplates",
        "ec2:DescribeNetworkAcls",
        "ec2:DescribeNetworkInterfaces",
        "ec2:DescribePlacementGroups",
```

```
        "ec2:DescribeRouteTables",
        "ec2:DescribeSecurityGroups",
        "ec2:DescribeSubnets",
        "ec2:DescribeVolumes",
        "ec2:DescribeVolumeStatus",
        "ec2:DescribeVpcAttribute",
        "ec2:DescribeVpcEndpoints",
        "ec2:DescribeVpcs"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "CreateDefaultSecurityGroupWithEMRTags",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateSecurityGroup"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:security-group/*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "CreateDefaultSecurityGroupInVPCWithEMRTags",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateSecurityGroup"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:vpc/*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
```

```
        "Sid": "TagOnCreateDefaultSecurityGroupWithEMRTags",
        "Effect": "Allow",
        "Action": [
          "ec2:CreateTags"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:security-group/*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true",
            "ec2:CreateAction": "CreateSecurityGroup"
          }
        }
      },
      {
        "Sid": "ManageSecurityGroups",
        "Effect": "Allow",
        "Action": [
          "ec2:AuthorizeSecurityGroupEgress",
          "ec2:AuthorizeSecurityGroupIngress",
          "ec2:RevokeSecurityGroupEgress",
          "ec2:RevokeSecurityGroupIngress"
        ],
        "Resource": [
          "*"
        ],
        "Condition": {
          "StringEquals": {
            "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
          }
        }
      },
      {
        "Sid": "CreateEMRPlacementGroups",
        "Effect": "Allow",
        "Action": [
          "ec2:CreatePlacementGroup"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:placement-group/pg-*"
        ]
      },
      {
```

```
        "Sid": "DeletePlacementGroups",
        "Effect": "Allow",
        "Action": [
          "ec2:DeletePlacementGroup"
        ],
        "Resource": [
          "*"
        ]
      },
      {
        "Sid": "AutoScaling",
        "Effect": "Allow",
        "Action": [
          "application-autoscaling:DeleteScalingPolicy",
          "application-autoscaling:DeregisterScalableTarget",
          "application-autoscaling:DescribeScalableTargets",
          "application-autoscaling:DescribeScalingPolicies",
          "application-autoscaling:PutScalingPolicy",
          "application-autoscaling:RegisterScalableTarget"
        ],
        "Resource": [
          "*"
        ]
      },
      {
        "Sid": "ResourceGroupsForCapacityReservations",
        "Effect": "Allow",
        "Action": [
          "resource-groups:ListGroupResources"
        ],
        "Resource": [
          "*"
        ]
      },
      {
        "Sid": "AutoScalingCloudWatch",
        "Effect": "Allow",
        "Action": [
          "cloudwatch:PutMetricAlarm",
          "cloudwatch:DeleteAlarms",
          "cloudwatch:DescribeAlarms"
        ],
        "Resource": [
          "arn:aws:cloudwatch:*:*:alarm:*_EMR_Auto_Scaling"
```

```
        ]
      },
      {
        "Sid": "PassRoleForAutoScaling",
        "Effect": "Allow",
        "Action": [
          "iam:PassRole"
        ],
        "Resource": [
          "arn:aws:iam::*:role/EMR_AutoScaling_DefaultRole"
        ],
        "Condition": {
          "StringLike": {
            "iam:PassedToService": "application-autoscaling.amazonaws.com*"
          }
        }
      },
      {
        "Sid": "PassRoleForEC2",
        "Effect": "Allow",
        "Action": [
          "iam:PassRole"
        ],
        "Resource": [
          "arn:aws:iam::*:role/EMR_EC2_DefaultRole"
        ],
        "Condition": {
          "StringLike": {
            "iam:PassedToService": "ec2.amazonaws.com*"
          }
        }
      },
      {
        "Sid": "CreateAndModifyEmrServiceVPCEndpoint",
        "Effect": "Allow",
        "Action": [
          "ec2:ModifyVpcEndpoint",
          "ec2:CreateVpcEndpoint"
        ],
        "Resource": [
          "arn:aws:ec2:*:*:vpc-endpoint/*",
          "arn:aws:ec2:*:*:subnet/*",
          "arn:aws:ec2:*:*:security-group/*",
          "arn:aws:ec2:*:*:vpc/*"
```

```
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "CreateEmrServiceVPCEndpoint",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateVpcEndpoint"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:vpc-endpoint/*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true",
          "aws:RequestTag/Name": "emr-service-vpce"
        }
      }
    },
    {
      "Sid": "TagEmrServiceVPCEndpoint",
      "Effect": "Allow",
      "Action": [
        "ec2:CreateTags"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:vpc-endpoint/*"
      ],
      "Condition": {
        "StringEquals": {
          "ec2:CreateAction": "CreateVpcEndpoint",
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true",
          "aws:RequestTag/Name": "emr-service-vpce"
        }
      }
    }
  ]
}
```

Su rol de servicio debe usar la siguiente política de confianza.

> **Important**
>
> La siguiente política de confianza incluye las claves de condición globales `aws:SourceArn` y `aws:SourceAccount` para limitar los permisos que concede a Amazon EMR a determinados recursos de su cuenta. Con estas, podrá protegerse contra el problema del suplente confuso.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowSTSAssumerole",
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/EMRServiceRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "123456789012"
        },
        "ArnLike": {
          "aws:SourceArn": "arn:aws:elasticmapreduce:*:123456789012:*"
        }
      }
    }
  ]
}
```

Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia)

La función de servicio para EC2 las instancias de clúster (también denominada perfil de EC2 instancia para Amazon EMR) es un tipo especial de función de servicio que se asigna a todas las instancias de un clúster de Amazon EMR cuando se lanza la EC2 instancia. Los procesos de

aplicación que se ejecutan sobre el ecosistema de Hadoop asumen este rol para los permisos, para interactuar así con otros productos de AWS .

Para obtener más información sobre las funciones de servicio para EC2 las instancias, consulte Uso de una función de IAM para conceder permisos a las aplicaciones que se ejecutan en EC2 instancias de Amazon en la Guía del usuario de IAM.

> **Important**
>
> El rol de servicio predeterminado para EC2 las instancias de clúster y su política administrada AWS predeterminada asociada `AmazonElasticMapReduceforEC2Role` están en vías de caducar, y no se proporcionan políticas AWS administradas sustitutivas. Tendrá que crear y especificar un perfil de instancia para reemplazar la política predeterminada y el rol obsoletos.

Política administrada y rol predeterminados

- El nombre del rol predeterminado es `EMR_EC2_DefaultRole`.
- La política administrada `EMR_EC2_DefaultRole` predeterminada, `AmazonElasticMapReduceforEC2Role`, está a punto de finalizar su soporte. En lugar de usar una política administrada predeterminada para el perfil de EC2 instancia, aplique políticas basadas en recursos a los buckets de S3 y otros recursos que Amazon EMR necesite, o use su propia política administrada por el cliente con un rol de IAM como perfil de instancia. Para obtener más información, consulte Crear un rol de servicio para las instancias de clúster con permisos de privilegios mínimos EC2 .

Lo siguiente muestra el contenido de la versión 3 de `AmazonElasticMapReduceforEC2Role`.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Resource": [
        "*"
```

```
        ],
        "Action": [
          "cloudwatch:*",
          "dynamodb:*",
          "ec2:Describe*",
          "elasticmapreduce:Describe*",
          "elasticmapreduce:ListBootstrapActions",
          "elasticmapreduce:ListClusters",
          "elasticmapreduce:ListInstanceGroups",
          "elasticmapreduce:ListInstances",
          "elasticmapreduce:ListSteps",
          "kinesis:CreateStream",
          "kinesis:DeleteStream",
          "kinesis:DescribeStream",
          "kinesis:GetRecords",
          "kinesis:GetShardIterator",
          "kinesis:MergeShards",
          "kinesis:PutRecord",
          "kinesis:SplitShard",
          "rds:Describe*",
          "s3:*",
          "sdb:*",
          "sns:*",
          "sqs:*",
          "glue:CreateDatabase",
          "glue:UpdateDatabase",
          "glue:DeleteDatabase",
          "glue:GetDatabase",
          "glue:GetDatabases",
          "glue:CreateTable",
          "glue:UpdateTable",
          "glue:DeleteTable",
          "glue:GetTable",
          "glue:GetTables",
          "glue:GetTableVersions",
          "glue:CreatePartition",
          "glue:BatchCreatePartition",
          "glue:UpdatePartition",
          "glue:DeletePartition",
          "glue:BatchDeletePartition",
          "glue:GetPartition",
          "glue:GetPartitions",
          "glue:BatchGetPartition",
          "glue:CreateUserDefinedFunction",
```

```
        "glue:UpdateUserDefinedFunction",
        "glue:DeleteUserDefinedFunction",
        "glue:GetUserDefinedFunction",
        "glue:GetUserDefinedFunctions"
      ],
      "Sid": "AllowCLOUDWATCH"
    }
  ]
}
```

Su rol de servicio debe usar la siguiente política de confianza.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "AllowSTSAssumerole",
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/EMR_EC2_DefaultRole"
    }
  ]
}
```

Crear un rol de servicio para las instancias de clúster con permisos de privilegios mínimos EC2

Como práctica recomendada, te recomendamos encarecidamente que crees un rol de servicio para las EC2 instancias del clúster y una política de permisos que tenga los permisos mínimos para otros AWS servicios que requiera tu aplicación.

La política administrada predeterminada, `AmazonElasticMapReduceforEC2Role`, proporciona los permisos que facilitan el lanzar un clúster inicial. Sin embargo, `AmazonElasticMapReduceforEC2Role` está en vías de quedar obsoleto y Amazon EMR no proporcionará una política predeterminada gestionada que AWS sustituya a la función obsoleta. Para

lanzar un clúster inicial, debe proporcionar una política basada en los recursos o en la identificación administrada por el cliente.

Las siguientes instrucciones de política facilitan ejemplos de permisos necesarios para las distintas características de Amazon EMR. Le recomendamos que utilice estos permisos para crear una política de permisos que restrinja el acceso tan solo a aquellas funciones y recursos que necesite el clúster. Todos los ejemplos de declaraciones de política utilizan la `us-west-2` región y el identificador de cuenta ficticio AWS . `123456789012` Sustituya estos según corresponda para el clúster.

Para obtener más información sobre la creación y la especificación de roles personalizados, consulte Personalización de roles de IAM con Amazon EMR.

> **Note**
>
> Si crea un rol de EMR personalizado para EC2, siga el flujo de trabajo básico, que crea automáticamente un perfil de instancia con el mismo nombre. Amazon EC2 le permite crear perfiles de instancia y roles con nombres diferentes, pero Amazon EMR no admite esta configuración y se produce un error de «perfil de instancia no válido» al crear el clúster.

Lectura y escritura de datos en Amazon S3 con EMRFS

Cuando una aplicación que se ejecuta en un clúster de Amazon EMR hace referencia a los datos mediante el `s3://mydata` formato, Amazon EMR utiliza el perfil de EC2 instancia para realizar la solicitud. Los clústeres suelen leer y escribir datos en Amazon S3 de esta forma, y Amazon EMR utiliza los permisos asociados a la función de servicio para las EC2 instancias de clúster de forma predeterminada. Para obtener más información, consulte Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3.

Dado que las funciones de IAM para EMRFS se basarán en los permisos asociados a la función de servicio para las EC2 instancias de clúster, como práctica recomendada, le recomendamos que utilice las funciones de IAM para EMRFS y que limite los permisos de EMRFS y Amazon S3 asociados a la función de servicio para las instancias de clúster. EC2

La siguiente instrucción de ejemplo señala los permisos que EMRFS necesita para hacer solicitudes a Amazon S3.

- *my-data-bucket-in-s3-for-emrfs-reads-and-writes* especifica el bucket de Amazon S3 en el que el clúster lee y escribe los datos y todas las subcarpetas con */*. Añada solo los buckets y carpetas que necesita su aplicación.
- La instrucción de política que permite realizar acciones de dynamodb solo es necesaria si la vista coherente de EMRFS está habilitada. *EmrFSMetadata* especifica la carpeta predeterminada para la vista coherente de EMRFS.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:AbortMultipartUpload",
        "s3:CreateBucket",
        "s3:DeleteObject",
        "s3:GetBucketVersioning",
        "s3:GetObject",
        "s3:GetObjectTagging",
        "s3:GetObjectVersion",
        "s3:ListBucket",
        "s3:ListBucketMultipartUploads",
        "s3:ListBucketVersions",
        "s3:ListMultipartUploadParts",
        "s3:PutBucketVersioning",
        "s3:PutObject",
        "s3:PutObjectTagging"
      ],
      "Resource": [
        "arn:aws:s3:::my-data-bucket-in-s3-for-emrfs-reads-and-writes",
        "arn:aws:s3:::my-data-bucket-in-s3-for-emrfs-reads-and-writes/*"
      ],
      "Sid": "AllowS3Abortmultipartupload"
    },
    {
      "Effect": "Allow",
      "Action": [
        "dynamodb:CreateTable",
        "dynamodb:BatchGetItem",
```

```
        "dynamodb:BatchWriteItem",
        "dynamodb:PutItem",
        "dynamodb:DescribeTable",
        "dynamodb:DeleteItem",
        "dynamodb:GetItem",
        "dynamodb:Scan",
        "dynamodb:Query",
        "dynamodb:UpdateItem",
        "dynamodb:DeleteTable",
        "dynamodb:UpdateTable"
      ],
      "Resource": [
        "arn:aws:dynamodb:*:123456789012:table/EmrFSMetadata"
      ],
      "Sid": "AllowDYNAMODBCreatetable"
    },
    {
      "Effect": "Allow",
      "Action": [
        "cloudwatch:PutMetricData",
        "dynamodb:ListTables",
        "s3:ListBucket"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowCLOUDWATCHPutmetricdata"
    },
    {
      "Effect": "Allow",
      "Action": [
        "sqs:GetQueueUrl",
        "sqs:ReceiveMessage",
        "sqs:DeleteQueue",
        "sqs:SendMessage",
        "sqs:CreateQueue"
      ],
      "Resource": [
        "arn:aws:sqs:*:123456789012:EMRFS-Inconsistency-*"
      ],
      "Sid": "AllowSQSGetqueueurl"
    }
  ]
```

```
}
```

Almacenamiento de archivos de registro en Amazon S3

La siguiente instrucción de política permite al clúster de Amazon EMR almacenar los archivos de registro en la ubicación de Amazon S3 indicada. En el ejemplo siguiente, cuando se creó el clúster, *s3://MyLoggingBucket/MyEMRClusterLogs* se especificó mediante la ubicación de la carpeta de registro S3 en la consola AWS CLI, mediante la `--log-uri` opción de o mediante el `LogUri` parámetro del `RunJobFlow` comando. Para obtener más información, consulte Archivar archivos de registro en Amazon S3.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::MyLoggingBucket/MyEMRClusterLogs/*"
      ],
      "Sid": "AllowS3Putobject"
    }
  ]
}
```

Uso del catálogo de datos de AWS Glue

La siguiente declaración de política permite las acciones que son necesarias si se utiliza el catálogo de datos de AWS Glue como almacén de aplicaciones. Para obtener más información, consulte Uso del catálogo de datos de AWS Glue como metaalmacén para Spark SQL, Uso del catálogo de datos de AWS Glue como metaalmacén para Hive y Uso de Presto con el catálogo de datos de AWS Glue en la Guía de versiones de Amazon EMR.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "glue:CreateDatabase",
        "glue:UpdateDatabase",
        "glue:DeleteDatabase",
        "glue:GetDatabase",
        "glue:GetDatabases",
        "glue:CreateTable",
        "glue:UpdateTable",
        "glue:DeleteTable",
        "glue:GetTable",
        "glue:GetTables",
        "glue:GetTableVersions",
        "glue:CreatePartition",
        "glue:BatchCreatePartition",
        "glue:UpdatePartition",
        "glue:DeletePartition",
        "glue:BatchDeletePartition",
        "glue:GetPartition",
        "glue:GetPartitions",
        "glue:BatchGetPartition",
        "glue:CreateUserDefinedFunction",
        "glue:UpdateUserDefinedFunction",
        "glue:DeleteUserDefinedFunction",
        "glue:GetUserDefinedFunction",
        "glue:GetUserDefinedFunctions"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowGLUECreatedatabase"
    }
  ]
}
```

Rol de servicio para el escalado automático en Amazon EMR (rol de Auto Scaling)

El rol de Auto Scaling de Amazon EMR sirve básicamente para lo mismo que el rol de servicio, pero permite acciones adicionales para entornos de escalado dinámico.

- El nombre del rol predeterminado es `EMR_AutoScaling_DefaultRole`.
- La política administrada predeterminada y asociada a `EMR_AutoScaling_DefaultRole` es `AmazonElasticMapReduceforAutoScalingRole`.

El contenido de la versión 1 de `AmazonElasticMapReduceforAutoScalingRole` se muestra a continuación.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "cloudwatch:DescribeAlarms",
        "elasticmapreduce:ListInstanceGroups",
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Sid": "AllowCLOUDWATCHDescribealarms"
    }
  ]
}
```

Su rol de servicio debe usar la siguiente política de confianza.

> **Important**
>
> La siguiente política de confianza incluye las claves de condición globales `aws:SourceArn` y `aws:SourceAccount` para limitar los permisos que concede a Amazon EMR a

determinados recursos de su cuenta. Con estas, podrá protegerse contra el problema del suplente confuso.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/ApplicationAutoScalingEMRRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "123456789012"
        },
        "ArnLike": {
          "aws:SourceArn": "arn:aws:application-
autoscaling:*:123456789012:scalable-target/*"
        }
      },
      "Sid": "AllowSTSAssumerole"
    }
  ]
}
```

Rol de servicio para Cuadernos de Amazon EMR

Cada cuaderno EMR necesita permisos para acceder a otros AWS recursos y realizar acciones. Las políticas de IAM asociadas a esta función de servicio proporcionan permisos para que el portátil interactúe con otros servicios. AWS Al crear un bloc de notas utilizando el AWS Management Console, se especifica un rol de AWS servicio. Puede utilizar el rol predeterminado, `EMR_Notebooks_DefaultRole` o especificar un rol que haya creado. Si no se ha creado un bloc de notas anteriormente, puede elegir crear el rol predeterminado.

- El nombre del rol predeterminado es `EMR_Notebooks_DefaultRole`.

- Las políticas administradas predeterminadas asociadas a `EMR_Notebooks_DefaultRole` son `AmazonElasticMapReduceEditorsRole` y `S3FullAccessPolicy`.

Su rol de servicio debe usar la siguiente política de confianza.

> **Important**
>
> La siguiente política de confianza incluye las claves de condición globales `aws:SourceArn` y `aws:SourceAccount` para limitar los permisos que concede a Amazon EMR a determinados recursos de su cuenta. Con estas, podrá protegerse contra el problema del suplente confuso.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/EMRServiceRole",
      "Condition": {
        "StringEquals": {
          "aws:SourceAccount": "123456789012"
        },
        "ArnLike": {
          "aws:SourceArn": "arn:aws:elasticmapreduce:*:123456789012:*"
        }
      },
      "Sid": "AllowSTSAssumerole"
    }
  ]
}
```

El contenido de la versión 1 de `AmazonElasticMapReduceEditorsRole` es el siguiente.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "ec2:AuthorizeSecurityGroupEgress",
        "ec2:AuthorizeSecurityGroupIngress",
        "ec2:CreateSecurityGroup",
        "ec2:DescribeSecurityGroups",
        "ec2:RevokeSecurityGroupEgress",
        "ec2:CreateNetworkInterface",
        "ec2:CreateNetworkInterfacePermission",
        "ec2:DeleteNetworkInterface",
        "ec2:DeleteNetworkInterfacePermission",
        "ec2:DescribeNetworkInterfaces",
        "ec2:ModifyNetworkInterfaceAttribute",
        "ec2:DescribeTags",
        "ec2:DescribeInstances",
        "ec2:DescribeSubnets",
        "ec2:DescribeVpcs",
        "elasticmapreduce:ListInstances",
        "elasticmapreduce:DescribeCluster",
        "elasticmapreduce:ListSteps"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Authorizesecuritygroupegress"
    },
    {
      "Effect": "Allow",
      "Action": [
        "ec2:CreateTags"
      ],
      "Resource": [
        "arn:aws:ec2:*:*:network-interface/*"
      ],
      "Condition": {
        "ForAllValues:StringEquals": {
          "aws:TagKeys": [
```

```
            "aws:elasticmapreduce:editor-id",
            "aws:elasticmapreduce:job-flow-id"
          ]
        }
      },
      "Sid": "AllowEC2Createtags"
    }
  ]
}
```

A continuación, se muestra el contenido de `S3FullAccessPolicy`. La `S3FullAccessPolicy` permite que su rol de servicio para Cuadernos de Amazon EMR realice todas las acciones de Amazon S3 en los objetos de su Cuenta de AWS. Al crear un rol de servicio personalizado para Cuadernos de Amazon EMR, debe otorgarle permisos a Amazon S3 para ese rol de servicio.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "s3:*"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowS3"
    }
  ]
}
```

Puede limitar el acceso de lectura y escritura para su rol de servicio a la ubicación de Amazon S3 en la que desee guardar los archivos de su cuaderno. Utilice el siguiente conjunto mínimo de permisos de Amazon S3.

```
"s3:PutObject",
"s3:GetObject",
```

```
"s3:GetEncryptionConfiguration",
"s3:ListBucket",
"s3:DeleteObject"
```

Si su bucket de Amazon S3 está cifrado, debe incluir los siguientes permisos para AWS Key Management Service.

```
"kms:Decrypt",
"kms:GenerateDataKey",
"kms:ReEncryptFrom",
"kms:ReEncryptTo",
"kms:DescribeKey"
```

Cuando vincula repositorios de Git a su cuaderno y necesita crear un secreto para el repositorio, debe agregar el permiso `secretsmanager:GetSecretValue` a la política de IAM asociada al rol de servicio de Cuadernos de Amazon EMR. A continuación se muestra una política de ejemplo:

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "VisualEditor0",
      "Effect": "Allow",
      "Action": [
        "secretsmanager:GetSecretValue"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Permisos del rol de servicio de Cuadernos de Amazon EMR

En esta tabla se enumeran las acciones que Cuadernos de Amazon EMR lleva a cabo con el rol de servicio, junto con los permisos necesarios para cada acción.

Acción	Permisos

Establezca un canal de red seguro entre un cuaderno y un clúster de Amazon EMR y lleve a cabo las acciones de limpieza necesarias.	`"ec2:CreateNetworkInterface",`
	`"ec2:CreateNetworkInterfacePermission",`
	`"ec2:DeleteNetworkInterface",`
	`"ec2:DeleteNetworkInterfacePermission",`
	`"ec2:DescribeNetworkInterfaces",`
	`"ec2:ModifyNetworkInterfaceAttribute",`
	`"ec2:AuthorizeSecurityGroupEgress",`
	`"ec2:AuthorizeSecurityGroupIngress",`
	`"ec2:CreateSecurityGroup",`
	`"ec2:DescribeSecurityGroups",`
	`"ec2:RevokeSecurityGroupEgress",`
	`"ec2:DescribeTags",`
	`"ec2:DescribeInstances",`
	`"ec2:DescribeSubnets",`
	`"ec2:DescribeVpcs",`
	`"elasticmapreduce:ListInstances",`
	`"elasticmapreduce:DescribeCluster",`
	`"elasticmapreduce:ListSteps"`

Usa las credenciales de Git almacenadas en AWS Secrets Manager para vincular los repositor ios de Git a un cuaderno.	`"secretsmanager:GetSecretValue"`
Aplique AWS etiquetas a la interfaz de red y a los grupos de seguridad predeterminados que EMR Notebooks crea al configurar el canal de red seguro. Para obtener más información, consulte Tagging AWS resources (Etiquetado de los recursos de).	`"ec2:CreateTags"`

Acceda a los archivos y metadatos de los cuadernos en Amazon S3 o cárguelos.	`"s3:PutObject",`
	`"s3:GetObject",`
	`"s3:GetEncryptionConfiguration",`
	`"s3:ListBucket",`

Acción	Permisos

	`"s3:DeleteObject"`
	
	Los siguientes permisos solo son necesarios si utiliza un bucket de Amazon S3 cifrado.
	
	`"kms:Decrypt",`
	`"kms:GenerateDataKey",`
	`"kms:ReEncryptFrom",`
	`"kms:ReEncryptTo",`
	`"kms:DescribeKey"`

Actualizaciones de EMR Notebooks a las políticas gestionadas AWS

Consulta los detalles sobre las actualizaciones de las políticas AWS gestionadas para los Notebooks EMR desde el 1 de marzo de 2021.

Cambio	Descripción	Fecha
`AmazonElasticMapReduceEditorsRole - Added permissions`	Cuadernos de Amazon EMR agregó los permisos `ec2:describeVPCs` y `elastmicmapreduce:ListSteps` a `AmazonElasticMapReduceEditorsRole` .	8 de febrero de 2023
Cuadernos de Amazon EMR comenzó a registrar cambios	EMR Notebooks comenzó a realizar un seguimiento de los cambios en sus políticas gestionadas AWS .	8 de febrero de 2023

Uso de roles vinculados a servicios para Amazon EMR

Amazon EMR utiliza funciones vinculadas a AWS Identity and Access Management servicios (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que se encuentra vinculado

directamente a Amazon EMR. Amazon EMR predefine las funciones vinculadas al servicio e incluyen todos los permisos que el servicio requiere para llamar a otros AWS servicios en su nombre.

Temas

- Uso de roles vinculados a servicios para Amazon EMR para limpieza
- Uso de roles vinculados a servicios con Amazon EMR para el registro de escritura anticipada

Para obtener información sobre otros servicios que admiten funciones vinculadas a servicios, consulte los AWS servicios que funcionan con IAM y busque los servicios con la palabra Sí en la columna Funciones vinculadas a servicios. Elija una opción Sí con un enlace para ver la documentación acerca del rol vinculado al servicio en cuestión.

Uso de roles vinculados a servicios para Amazon EMR para limpieza

Amazon EMR utiliza funciones vinculadas a AWS Identity and Access Management servicios (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que se encuentra vinculado directamente a Amazon EMR. Amazon EMR predefine las funciones vinculadas al servicio e incluyen todos los permisos que el servicio requiere para llamar a otros AWS servicios en su nombre.

Los roles vinculados al servicio funcionan junto con el rol de servicio de Amazon EMR y el EC2 perfil de instancia de Amazon para Amazon EMR. Para obtener más información acerca del rol de servicio y del perfil de instancia, consulte Configuración de los roles de servicio de IAM de los permisos de Amazon EMR para los servicios y recursos de AWS.

Un rol vinculado al servicio simplifica la configuración de Amazon EMR porque ya no tendrá que agregar manualmente los permisos requeridos. Amazon EMR define los permisos de sus roles vinculados a servicios y, a menos que esté definido de otra manera, solo Amazon EMR puede asumir sus roles. Los permisos definidos incluyen las políticas de confianza y de permisos, y que la política de permisos no se pueda asociar a ninguna otra entidad de IAM.

Puede eliminar este rol vinculado a servicios para Amazon EMR solo después de eliminar cualquier recurso relacionado y finalizar todos los clústeres de EMR de la cuenta. Esto protege sus recursos de Amazon EMR, puesto que se evita que se puedan eliminar accidentalmente permisos de acceso a los recursos.

Uso de roles vinculados a servicios para limpieza

Amazon EMR utiliza la AWSServiceRoleForEMRCleanupfunción basada en servicios para conceder a Amazon EMR permiso para cancelar y eliminar los EC2 recursos de Amazon en su nombre si la

función vinculada al servicio de Amazon EMR pierde esa capacidad. Si aún no existe, Amazon EMR crea el rol vinculado a servicios automáticamente durante la creación del clúster.

El rol AWSService RoleFor EMRCleanup vinculado al servicio confía en que los siguientes servicios asuman el rol:

- `elasticmapreduce.amazonaws.com`

La política de permisos de roles AWSService RoleFor EMRCleanup vinculados al servicio permite a Amazon EMR realizar las siguientes acciones en los recursos especificados:

- Acción: `DescribeInstances` en `ec2`
- Acción: `DescribeLaunchTemplates` en `ec2`
- Acción: `DeleteLaunchTemplate` en `ec2`
- Acción: `DescribeSpotInstanceRequests` en `ec2`
- Acción: `ModifyInstanceAttribute` en `ec2`
- Acción: `TerminateInstances` en `ec2`
- Acción: `CancelSpotInstanceRequests` en `ec2`
- Acción: `DeleteNetworkInterface` en `ec2`
- Acción: `DescribeInstanceAttribute` en `ec2`
- Acción: `DescribeVolumeStatus` en `ec2`
- Acción: `DescribeVolumes` en `ec2`
- Acción: `DetachVolume` en `ec2`
- Acción: `DeleteVolume` en `ec2`
- Acción: `DescribePlacementGroups` en `ec2`
- Acción: `DeletePlacementGroup` en `ec2`

Debe configurar permisos para permitir a una entidad de IAM (como un usuario, grupo o rol) crear, editar o eliminar un rol vinculado a servicios.

Creación de un rol vinculado a un servicio para Amazon EMR

No necesita crear el rol manualmente. AWSService RoleFor EMRCleanup Cuando lanza un clúster, ya sea por primera vez o cuando el rol AWSService RoleFor EMRCleanup vinculado al servicio no está presente, Amazon EMR crea el rol vinculado al AWSService RoleFor EMRCleanup servicio

automáticamente. Debe tener permisos para crear un rol vinculado a servicios. Para ver un ejemplo de declaración que añade esta capacidad a la política de permisos de una entidad de IAM (como un usuario, un grupo o un rol):

Agregue la siguiente instrucción a la política de permisos de la entidad de IAM que tiene que crear el rol vinculado al servicio:

```
{
   "Effect": "Allow",
   "Action": [
       "iam:CreateServiceLinkedRole"
   ],
   "Resource": "arn:aws:iam::*:role/aws-service-role/elasticmapreduce.amazonaws.com*/
AWSServiceRoleForEMRCleanup*",
   "Condition": {"StringLike": {"iam:AWSServiceName": "ops.emr-
serverless.amazonaws.com"}}
 }
```

> **Important**
>
> Si utilizaste Amazon EMR antes del 24 de octubre de 2017, cuando no se admitían las funciones vinculadas a servicios, Amazon EMR creó la AWSService RoleFor EMRCleanup función vinculada a servicios en tu cuenta. Para obtener más información, consulte Un nuevo rol ha aparecido en la cuenta de IAM.

Edición de un rol vinculado a un servicio para Amazon EMR

Amazon EMR no le permite editar el rol vinculado al AWSService RoleFor EMRCleanup servicio. Después de crear un rol vinculado a servicios, no puede cambiarle el nombre a este rol vinculado a servicios, ya que varias entidades pueden hacer referencia al rol vinculado a servicios. No obstante, puede editar la descripción del rol vinculado a servicios mediante IAM.

Edición de la descripción de un rol vinculado a un servicio (consola de IAM)

Puede utilizar la consola de IAM para editar la descripción de un rol vinculado a un servicio.

Para editar la descripción de un rol vinculado a un servicio (consola)

1. En el panel de navegación de la consola de IAM, elija Roles.

2. Seleccione el nombre del rol que desea modificar.
3. En el extremo derecho de Descripción del rol, seleccione Editar.
4. Ingrese una descripción nueva en el cuadro y elija Save changes (Guardar cambios).

Edición de la descripción de un rol vinculado a un servicio (CLI de IAM)

Puede utilizar los comandos de IAM del AWS Command Line Interface para editar la descripción de un rol vinculado a un servicio.

Para cambiar la descripción de un rol vinculado a un servicio (CLI)

1. (Opcional) Para ver la descripción actual de un rol, ejecute uno de los siguientes comandos:

   ```
   $ aws iam get-role --role-name role-name
   ```

   Utilice el nombre del rol, no el ARN, para hacer referencia a los roles con los comandos de CLI. Por ejemplo, si un rol tiene el ARN `arn:aws:iam::123456789012:role/myrole`, debe referirse a él como **`myrole`**.
2. Para actualizar la descripción de un rol vinculado a un servicio, ejecute uno de los siguientes comandos:

   ```
   $ aws iam update-role-description --role-name role-name --description description
   ```

Edición de la descripción de un rol vinculado a un servicio (API de IAM)

Puede utilizar la API de IAM para editar la descripción de un rol vinculado a un servicio.

Para cambiar la descripción de un rol vinculado a un servicio (API)

1. (Opcional) Para ver la descripción actual de una función, ejecute el siguiente comando:

   API de IAM: GetRole
2. Para actualizar la descripción de una función, use el siguiente comando:

   API de IAM: UpdateRoleDescription

Eliminación de un rol vinculado a un servicio para Amazon EMR

Si ya no necesita usar una característica o servicio que requieran un rol vinculado a servicios, le recomendamos que elimine dicho rol vinculado a servicios. De esta forma, no tendrá una entidad no utilizada cuya supervisión o mantenimiento no se realizan de forma activa. Sin embargo, debe limpiar el rol vinculado al servicio antes de eliminarlo.

Saneamiento de un rol vinculado a servicios

Antes de poder utilizar IAM para eliminar un rol vinculado a servicios, primero debe confirmar que dicho rol vinculado a servicios no tiene sesiones activas y eliminar los recursos que utiliza el rol vinculado a servicios.

Para comprobar si el rol vinculado a un servicio tiene una sesión activa en la consola de IAM

1. Abra la consola de IAM en https://console.aws.amazon.com/iam/.
2. Seleccione Roles en el panel de navegación. Seleccione el nombre (no la casilla de verificación) del rol vinculado al AWSService RoleFor EMRCleanup servicio.
3. En la página Resumen del rol vinculado a servicios seleccionado, elija Asesor de acceso.
4. En la pestaña Access Advisor, revise la actividad reciente del rol vinculado al servicio.

> **Note**
>
> Si no está seguro de si Amazon EMR utiliza el rol AWSService RoleFor EMRCleanup vinculado al servicio, puede intentar eliminar el rol vinculado al servicio. Si el servicio está utilizando el rol vinculado a servicios, este no podrá eliminarse y podrá ver las regiones en las que se está utilizando el rol vinculado a servicios. Si el rol vinculado a servicios se está utilizando, debe esperar a que la sesión finalice para poder eliminar el rol vinculado a servicios. No se puede revocar la sesión de un rol vinculado a servicios.

Para eliminar los recursos de Amazon EMR utilizados por el AWSService RoleFor EMRCleanup

- Termine todos los clústeres de su cuenta. Para obtener más información, consulte Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera..

Eliminación de un rol vinculado a un servicio (consola de IAM)

Puede utilizar la consola de IAM para eliminar un rol vinculado a un servicio.

Para eliminar un rol vinculado a un servicio (consola)

1. Abra la consola de IAM en https://console.aws.amazon.com/iam/.
2. Seleccione Roles en el panel de navegación. Seleccione la casilla de verificación situada junto a ella AWSService RoleForEMRCleanup, no el nombre o la fila en sí.
3. En Role actions (Acciones de rol) en la parte superior de la página, elija Delete role (Eliminar rol).
4. En el cuadro de diálogo de confirmación, revise los datos del servicio al que se accedió por última vez, que muestran cuándo accedió por última vez a un AWS servicio cada uno de los roles seleccionados. Esto lo ayuda a confirmar si el rol está actualmente activo. Para continuar, elija Yes, Delete.
5. Consulte las notificaciones de la consola de IAM para supervisar el progreso de la eliminación del rol vinculado al servicio. Como el proceso de eliminación del rol vinculado a servicios de IAM es asíncrono, dicha tarea puede realizarse correctamente o fallar después de que envía la solicitud de eliminación del rol vinculado a servicios. Si la tarea no se realiza correctamente, puede seleccionar View details (Ver detalles) o View Resources (Ver recursos) desde las notificaciones para obtener información sobre el motivo por el que no se pudo eliminar el rol. Si la eliminación no pudo producirse porque hay recursos en el servicio que está utilizando el rol, entonces el motivo del error incluye una lista de recursos.

Eliminación de un rol vinculado a un servicio (CLI de IAM)

Puede utilizar los comandos de IAM desde el AWS Command Line Interface para eliminar un rol vinculado a un servicio. Como los roles vinculados a servicios no se puede eliminar si están en uso o tienen recursos asociados, debe enviar una solicitud de eliminación. Si estas condiciones no se cumplen, dicha solicitud se puede denegar.

Para eliminar un rol vinculado a un servicio (CLI)

1. Para comprobar el estado de la tarea de eliminación, debe apuntar el valor de `deletion-task-id` de la respuesta. Escriba el siguiente comando para enviar una solicitud de eliminación de un rol vinculado a un servicio:

```
$ aws iam delete-service-linked-role --role-name AWSServiceRoleForEMRCleanup
```

2. Escriba el siguiente comando para comprobar el estado de la tarea de eliminación:

```
$ aws iam get-service-linked-role-deletion-status --deletion-task-id deletion-task-id
```

El estado de la tarea de eliminación puede ser `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED` o `FAILED`. Si ocurre un error durante la eliminación, la llamada devuelve el motivo del error para que pueda resolver el problema.

Eliminación de un rol vinculado a un servicio (API de IAM)

Puede utilizar la API de IAM para eliminar un rol vinculado a un servicio. Como los roles vinculados a servicios no se puede eliminar si están en uso o tienen recursos asociados, debe enviar una solicitud de eliminación. Si estas condiciones no se cumplen, dicha solicitud se puede denegar.

Para eliminar un rol vinculado a un servicio (API)

1. Para enviar una solicitud de eliminación de un rol vinculado a un servicio, llama. DeleteServiceLinkedRole En la solicitud, especifique el nombre del AWSService RoleFor EMRCleanup rol.

   Para comprobar el estado de la tarea de eliminación, debe apuntar el valor de `DeletionTaskId` de la respuesta.
2. Para comprobar el estado de la tarea de eliminación, realice una llamada a GetServiceLinkedRoleDeletionStatus. En la solicitud, especifique el valor de `DeletionTaskId`.

   El estado de la tarea de eliminación puede ser `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED` o `FAILED`. Si ocurre un error durante la eliminación, la llamada devuelve el motivo del error para que pueda resolver el problema.

Regiones compatibles para AWSService RoleFor EMRCleanup

Amazon EMR admite el uso de la función AWSService RoleFor EMRCleanup vinculada al servicio en las siguientes regiones.

Nombre de la región	Identidad de la región	Compatibilidad en Amazon EMR
Este de EE. UU. (Norte de Virginia)	us-east-1	Sí

Nombre de la región	Identidad de la región	Compatibilidad en Amazon EMR
Este de EE. UU. (Ohio)	us-east-2	Sí
Oeste de EE. UU. (Norte de California)	us-west-1	Sí
Oeste de EE. UU. (Oregón)	us-west-2	Sí
Asia-Pacífico (Bombay)	ap-south-1	Sí
Asia Pacífico (Osaka)	ap-northeast-3	Sí
Asia-Pacífico (Seúl)	ap-northeast-2	Sí
Asia-Pacífico (Singapur)	ap-southeast-1	Sí
Asia-Pacífico (Sídney)	ap-southeast-2	Sí
Asia-Pacífico (Tokio)	ap-northeast-1	Sí
Canadá (centro)	ca-central-1	Sí
Europa (Fráncfort)	eu-central-1	Sí
Europa (Irlanda)	eu-west-1	Sí
Europa (Londres)	eu-west-2	Sí
Europa (París)	eu-west-3	Sí
América del Sur (São Paulo)	sa-east-1	Sí

Uso de roles vinculados a servicios con Amazon EMR para el registro de escritura anticipada

Amazon EMR utiliza funciones vinculadas a AWS Identity and Access Management servicios (IAM). Un rol vinculado a un servicio es un tipo único de rol de IAM que se encuentra vinculado directamente a Amazon EMR. Amazon EMR predefine las funciones vinculadas al servicio e incluyen todos los permisos que el servicio requiere para llamar a otros AWS servicios en su nombre.

Los roles vinculados al servicio funcionan junto con el rol de servicio de Amazon EMR y el EC2 perfil de instancia de Amazon para Amazon EMR. Para obtener más información acerca del rol de servicio y del perfil de instancia, consulte Configuración de los roles de servicio de IAM de los permisos de Amazon EMR para los servicios y recursos de AWS.

Un rol vinculado al servicio simplifica la configuración de Amazon EMR porque ya no tendrá que agregar manualmente los permisos requeridos. Amazon EMR define los permisos de sus roles vinculados a servicios y, a menos que esté definido de otra manera, solo Amazon EMR puede asumir sus roles. Los permisos definidos incluyen las políticas de confianza y de permisos, y que la política de permisos no se pueda asociar a ninguna otra entidad de IAM.

Puede eliminar este rol vinculado a servicios para Amazon EMR solo después de eliminar los recursos relacionados y finalizar todos los clústeres de EMR de la cuenta. Esto protege sus recursos de Amazon EMR, puesto que se evita que se puedan eliminar accidentalmente permisos de acceso a los recursos.

Permisos de roles vinculados a servicios para registro de escritura anticipada (WAL)

Amazon EMR utiliza la función AWSServiceRoleForEMRWAL vinculada al servicio para recuperar el estado de un clúster.

La función vinculada al servicio de AWSService RoleFor EMRWAL confía en los siguientes servicios para que la asuman:

- `emrwal.amazonaws.com`

La política de permisos de `EMRDescribeClusterPolicyForEMRWAL` para el rol vinculado a servicios permite a Amazon EMR que complete las siguientes acciones en los recursos especificados:

- Acción: `DescribeCluster` en `*`

Debe configurar los permisos para permitir que una entidad de IAM (en este caso, Amazon EMR WAL) cree, edite o elimine un rol vinculado a servicios. Añada las siguientes instrucciones, según sea necesario, a la política de permisos de su perfil de instancia:

CreateServiceLinkedRole

Para permitir que una entidad de IAM cree la función vinculada al servicio de EMRWAL AWSService RoleFor

Agregue la siguiente instrucción a la política de permisos de la entidad de IAM que tiene que crear el rol vinculado al servicio:

```
{
    "Effect": "Allow",
    "Action": [
        "iam:CreateServiceLinkedRole",
        "iam:PutRolePolicy"
    ],
    "Resource": "arn:aws:iam::*:role/aws-service-role/emrwal.amazonaws.com*/
AWSServiceRoleForEMRWAL*",
    "Condition": {
        "StringLike": {
            "iam:AWSServiceName": [
                "emrwal.amazonaws.com",
                "elasticmapreduce.amazonaws.com.cn"
            ]
        }
    }
}
```

UpdateRoleDescription

Para permitir que una entidad de IAM edite la descripción de la función vinculada al servicio de EMRWAL AWSService RoleFor

Agregue la siguiente instrucción a la política de permisos de la entidad de IAM que tiene que editar la descripción del rol vinculado al servicio:

```
{
    "Effect": "Allow",
    "Action": [
        "iam:UpdateRoleDescription"
    ],
    "Resource": "arn:aws:iam::*:role/aws-service-role/emrwal.amazonaws.com*/
AWSServiceRoleForEMRWAL*",
    "Condition": {
        "StringLike": {
            "iam:AWSServiceName": [
                "emrwal.amazonaws.com",
                "elasticmapreduce.amazonaws.com.cn"
            ]
```

```
        }
    }
}
```

DeleteServiceLinkedRole

Para permitir que una entidad de IAM elimine la función vinculada al servicio de EMRWAL AWSService RoleFor

Agregue la siguiente instrucción a la política de permisos de la entidad de IAM que tiene que eliminar un rol vinculado al servicio:

```
{
    "Effect": "Allow",
    "Action": [
        "iam:DeleteServiceLinkedRole",
        "iam:GetServiceLinkedRoleDeletionStatus"
    ],
    "Resource": "arn:aws:iam::*:role/aws-service-role/elasticmapreduce.amazonaws.com*/
AWSServiceRoleForEMRCleanup*",
    "Condition": {
        "StringLike": {
            "iam:AWSServiceName": [
                "emrwal.amazonaws.com",
                "elasticmapreduce.amazonaws.com.cn"
            ]
        }
    }
}
```

Creación de un rol vinculado a un servicio para Amazon EMR

No es necesario crear manualmente el rol EMRWAL. AWSService RoleFor Amazon EMR crea este rol vinculado al servicio automáticamente cuando crea un espacio de trabajo de WAL con la CLI de EMRWAL o desde AWS CloudFormation, o HBase creará el rol vinculado al servicio cuando configure un espacio de trabajo para Amazon EMR WAL y el rol vinculado al servicio aún no existe. Debe tener permisos para crear un rol vinculado a servicios. Para instrucciones de ejemplo que agreguen esta capacidad a la política de permisos de una entidad de IAM (como un usuario, grupo o rol), consulte la sección anterior, Permisos de roles vinculados a servicios para registro de escritura anticipada (WAL).

Edición de un rol vinculado a un servicio para Amazon EMR

Amazon EMR no le permite editar el rol vinculado al servicio AWSService RoleFor EMRWAL. Después de crear un rol vinculado a servicios, no puede cambiarle el nombre a este rol vinculado a servicios, ya que varias entidades pueden hacer referencia al rol vinculado a servicios. No obstante, puede editar la descripción del rol vinculado a servicios mediante IAM.

Edición de la descripción de un rol vinculado a un servicio (consola de IAM)

Puede utilizar la consola de IAM para editar la descripción de un rol vinculado a un servicio.

Para editar la descripción de un rol vinculado a un servicio (consola)

1. En el panel de navegación de la consola de IAM, elija Roles.
2. Seleccione el nombre del rol que desea modificar.
3. En el extremo derecho de Descripción del rol, seleccione Editar.
4. Ingrese una descripción nueva en el cuadro y elija Save changes (Guardar cambios).

Edición de la descripción de un rol vinculado a un servicio (CLI de IAM)

Puede utilizar los comandos de IAM del AWS Command Line Interface para editar la descripción de un rol vinculado a un servicio.

Para cambiar la descripción de un rol vinculado a un servicio (CLI)

1. (Opcional) Para ver la descripción actual de un rol, ejecute uno de los siguientes comandos:

   ```
   $ aws iam get-role --role-name role-name
   ```

   Utilice el nombre del rol, no el ARN, para hacer referencia a los roles con los comandos de CLI. Por ejemplo, si un rol tiene el ARN `arn:aws:iam::123456789012:role/myrole`, debe referirse a él como **`myrole`**.
2. Para actualizar la descripción de un rol vinculado a un servicio, ejecute uno de los siguientes comandos:

   ```
   $ aws iam update-role-description --role-name role-name --description description
   ```

Edición de la descripción de un rol vinculado a un servicio (API de IAM)

Puede utilizar la API de IAM para editar la descripción de un rol vinculado a un servicio.

Para cambiar la descripción de un rol vinculado a un servicio (API)

1. (Opcional) Para ver la descripción actual de una función, ejecute el siguiente comando:

   API de IAM: GetRole
2. Para actualizar la descripción de una función, use el siguiente comando:

   API de IAM: UpdateRoleDescription

Eliminación de un rol vinculado a un servicio para Amazon EMR

Si ya no necesita usar una característica o servicio que requieran un rol vinculado a servicios, le recomendamos que elimine dicho rol vinculado a servicios. De esta forma, no tendrá una entidad no utilizada cuya supervisión o mantenimiento no se realizan de forma activa. Sin embargo, debe limpiar el rol vinculado al servicio antes de eliminarlo.

> **Note**
>
> La operación de registro de escritura anticipada no se ve afectada si eliminas la función AWSService RoleFor EMRWAL, pero Amazon EMR no eliminará automáticamente los registros que creó una vez que el clúster de EMR finalice. Por lo tanto, tendrá que eliminar manualmente los registros de Amazon EMR WAL si elimina el rol vinculado a servicios.

Limpiar un rol vinculado a servicios

Antes de poder utilizar IAM para eliminar un rol vinculado a un servicio, primero debe confirmar que dicho rol no tiene sesiones activas y eliminar los recursos que utiliza.

Para comprobar si el rol vinculado a un servicio tiene una sesión activa en la consola de IAM

1. Abra la consola de IAM en https://console.aws.amazon.com/iam/.
2. Seleccione Roles en el panel de navegación. Seleccione el nombre (no la casilla de verificación) del rol de EMRWAL. AWSService RoleFor
3. En la página Summary (Resumen) del rol seleccionado, seleccione Access Advisor.

4. En la pestaña Access Advisor, revise la actividad reciente del rol vinculado al servicio.

> **Note**
>
> Si no está seguro de si Amazon EMR utiliza la función AWSService RoleFor EMRWAL, puede intentar eliminar la función vinculada al servicio. Si el servicio está utilizando el rol, este no podrá eliminarse y podrá ver las regiones en las que se está utilizando el rol vinculado a servicios. Si el rol vinculado a servicios se está utilizando, debe esperar a que la sesión finalice para poder eliminar el rol vinculado a servicios. No se puede revocar la sesión de un rol vinculado a servicios.

Para eliminar los recursos de Amazon EMR utilizados por el EMRWAL AWSService RoleFor

- Termine todos los clústeres de su cuenta. Para obtener más información, consulte Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera..

Eliminación de un rol vinculado a un servicio (consola de IAM)

Puede utilizar la consola de IAM para eliminar un rol vinculado a un servicio.

Para eliminar un rol vinculado a un servicio (consola)

1. Abra la consola de IAM en https://console.aws.amazon.com/iam/.
2. Seleccione Roles en el panel de navegación. Seleccione la casilla de verificación situada junto a AWSService RoleFor EMRWAL, no el nombre o la fila en sí.
3. En Role actions (Acciones de rol) en la parte superior de la página, elija Delete role (Eliminar rol).
4. En el cuadro de diálogo de confirmación, revise los datos del servicio al que se accedió por última vez, que muestran cuándo accedió por última vez a un AWS servicio cada uno de los roles seleccionados. Esto lo ayuda a confirmar si el rol está actualmente activo. Para continuar, elija Yes, Delete.
5. Consulte las notificaciones de la consola de IAM para supervisar el progreso de la eliminación del rol vinculado al servicio. Como el proceso de eliminación del rol vinculado al servicio de IAM es asíncrono, dicha tarea puede realizarse correctamente o fallar después de que envía la solicitud de eliminación. Si la tarea no se realiza correctamente, puede seleccionar View details (Ver detalles) o View Resources (Ver recursos) desde las notificaciones para obtener

información sobre el motivo por el que no se pudo eliminar el rol. Si la eliminación no pudo producirse porque hay recursos en el servicio que está utilizando el rol, entonces el motivo del error incluye una lista de recursos.

Eliminación de un rol vinculado a un servicio (CLI de IAM)

Puede utilizar los comandos de IAM desde el AWS Command Line Interface para eliminar un rol vinculado a un servicio. Como los roles vinculados a servicios no se puede eliminar si están en uso o tienen recursos asociados, debe enviar una solicitud de eliminación. Si estas condiciones no se cumplen, dicha solicitud se puede denegar.

Para eliminar un rol vinculado a un servicio (CLI)

1. Para comprobar el estado de la tarea de eliminación, debe apuntar el valor de `deletion-task-id` de la respuesta. Escriba el siguiente comando para enviar una solicitud de eliminación de un rol vinculado a un servicio:

   ```
   $ aws iam delete-service-linked-role --role-name AWSServiceRoleForEMRWAL
   ```

2. Escriba el siguiente comando para comprobar el estado de la tarea de eliminación:

   ```
   $ aws iam get-service-linked-role-deletion-status --deletion-task-id deletion-task-id
   ```

   El estado de la tarea de eliminación puede ser `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED` o `FAILED`. Si ocurre un error durante la eliminación, la llamada devuelve el motivo del error para que pueda resolver el problema.

Eliminación de un rol vinculado a un servicio (API de IAM)

Puede utilizar la API de IAM para eliminar un rol vinculado a un servicio. Como los roles vinculados a servicios no se puede eliminar si están en uso o tienen recursos asociados, debe enviar una solicitud de eliminación. Si estas condiciones no se cumplen, dicha solicitud se puede denegar.

Para eliminar un rol vinculado a un servicio (API)

1. Para enviar una solicitud de eliminación de un rol vinculado a un servicio, llama. DeleteServiceLinkedRole En la solicitud, especifique el nombre del rol de AWSService RoleFor EMRWAL.

Para comprobar el estado de la tarea de eliminación, debe apuntar el valor de `DeletionTaskId` de la respuesta.

2. Para comprobar el estado de la tarea de eliminación, realice una llamada a GetServiceLinkedRoleDeletionStatus. En la solicitud, especifique el valor de `DeletionTaskId`.

   El estado de la tarea de eliminación puede ser `NOT_STARTED`, `IN_PROGRESS`, `SUCCEEDED` o `FAILED`. Si ocurre un error durante la eliminación, la llamada devuelve el motivo del error para que pueda resolver el problema.

Regiones compatibles con EMRWAL AWSService RoleFor

Amazon EMR admite el uso de la función vinculada AWSService RoleFor al servicio EMRWAL en las siguientes regiones.

Nombre de la región	Identidad de la región	Compatibilidad en Amazon EMR
Este de EE. UU. (Norte de Virginia)	us-east-1	Sí
Este de EE. UU. (Ohio)	us-east-2	Sí
Oeste de EE. UU. (Norte de California)	us-west-1	Sí
Oeste de EE. UU. (Oregón)	us-west-2	Sí
Asia-Pacífico (Bombay)	ap-south-1	Sí
Asia-Pacífico (Singapur)	ap-southeast-1	Sí
Asia-Pacífico (Sídney)	ap-southeast-2	Sí
Asia-Pacífico (Tokio)	ap-northeast-1	Sí
Europa (Fráncfort)	eu-central-1	Sí
Europa (Irlanda)	eu-west-1	Sí

Personalización de roles de IAM con Amazon EMR

Es posible que quiera personalizar los permisos y los roles de servicio de IAM para limitar los privilegios de acuerdo con los requisitos de seguridad. Para personalizar los permisos, le recomendamos que cree nuevos roles y políticas. Comience con los permisos de las políticas administradas para los roles predeterminados (por ejemplo, `AmazonElasticMapReduceforEC2Role` y `AmazonElasticMapReduceRole`). A continuación, copie y pegue el contenido de las nuevas instrucciones de política, modifique los permisos según corresponda y asocie las políticas de permisos modificadas a los roles que cree. Debe disponer de los permisos de IAM adecuados para trabajar con los roles y las políticas. Para obtener más información, consulte Cómo permitir a los usuarios y grupos crear y modificar roles.

Si crea un rol de EMR personalizado para EC2, siga el flujo de trabajo básico, que crea automáticamente un perfil de instancia con el mismo nombre. Amazon EC2 le permite crear perfiles de instancia y roles con nombres diferentes, pero Amazon EMR no admite esta configuración y se produce un error de «perfil de instancia no válido» al crear el clúster.

> **Important**
>
> Las políticas insertadas no se actualizan automáticamente cuando cambian los requisitos de servicio. Si crea y adjunta políticas insertadas, tenga en cuenta que se pueden producir actualizaciones de servicio que provoquen errores de permisos de forma repentina. Para obtener más información, consulte Políticas administradas y políticas insertadas en la Guía del usuario de IAM y Cómo especificar roles de IAM personalizados al crear un clúster.

Para obtener más información sobre cómo trabajar con los roles de IAM, consulte los siguientes temas en la Guía del usuario de IAM:

- Crear un rol para delegar permisos a un servicio AWS
- Modificación de un rol
- Eliminación de un rol

Cómo especificar roles de IAM personalizados al crear un clúster

Al crear un clúster, debe especificar el rol de servicio para Amazon EMR y el rol para el perfil de EC2 instancia de Amazon. El usuario que crea los clústeres necesita permisos para recuperar y asignar

funciones a Amazon EMR y EC2 a las instancias. De lo contrario, se producirá un EC2 error en la cuenta que no está autorizada a realizar llamadas. Para obtener más información, consulte Cómo permitir a los usuarios y grupos crear y modificar roles.

Uso de la consola para especificar roles personalizados

Al crear un clúster, puede especificar un rol de servicio personalizado para Amazon EMR, un rol personalizado para el perfil de EC2 instancia y un rol de Auto Scaling personalizado mediante las opciones avanzadas. Al utilizar las opciones rápidas, se especifican el rol de servicio predeterminado y el rol predeterminado para el perfil de la EC2 instancia. Para obtener más información, consulte Roles de servicio de IAM utilizados por Amazon EMR.

Console

Especificar roles de IAM personalizados mediante la consola

Al crear un clúster con la consola, debe especificar un rol de servicio personalizado para Amazon EMR y un rol personalizado para el perfil de la EC2 instancia. Para obtener más información, consulte Roles de servicio de IAM utilizados por Amazon EMR.

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Configuración y permisos de seguridad, busque los campos Rol de IAM para el perfil de instancia y Rol de servicio para Amazon EMR. Seleccione un rol en la lista para cada tipo de rol. Solo aparecerán los roles de la cuenta que tengan la política de confianza adecuada para ese tipo de rol.
4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

Utilícela AWS CLI para especificar funciones personalizadas

Puede especificar un rol de servicio para Amazon EMR y un rol de servicio para las EC2 instancias de clúster de forma explícita mediante las opciones con el `create-cluster` comando de. AWS CLI Utilice la opción `--service-role` para especificar el rol de servicio. Utilice el `InstanceProfile` argumento de la `--ec2-attributes` opción para especificar el rol del perfil de EC2 instancia.

El rol de Auto Scaling se especifica mediante la opción `--auto-scaling-role`. Para obtener más información, consulte Uso del escalado automático con una política personalizada para grupos de instancias en Amazon EMR.

Para especificar funciones de IAM personalizadas mediante el AWS CLI

- El siguiente comando especifica la función de servicio personalizada y una función personalizada para el perfil de la EC2 instancia al lanzar un clúster. *MyCustomServiceRoleForEMR* *MyCustomServiceRoleForClusterEC2Instances* Este ejemplo utiliza el rol de Amazon EMR predeterminado.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "Test cluster" --release-label emr-7.10.0 \
--applications Name=Hive Name=Pig --service-role MyCustomServiceRoleForEMR \
--ec2-attributes InstanceProfile=MyCustomServiceRoleForClusterEC2Instances,\
KeyName=myKey --instance-type m5.xlarge --instance-count 3
```

Puede utilizar estas opciones para especificar roles predeterminados de forma explícita en lugar de usar la opción `--use-default-roles`. La `--use-default-roles` opción especifica la función de servicio y la función del perfil de EC2 instancia definido en el `config` archivo de AWS CLI.

El siguiente ejemplo muestra el contenido de un `config` archivo para AWS CLI las funciones personalizadas especificadas para Amazon EMR. Con este archivo de configuración, cuando se especifica la opción `--use-default-roles`, el clúster se crea con *MyCustomServiceRoleForEMR* y *MyCustomServiceRoleForClusterEC2Instances*. De forma predeterminada, el archivo `config` especifica el valor predeterminado `service_role` como `AmazonElasticMapReduceRole` y el valor predeterminado `instance_profile` como `EMR_EC2_DefaultRole`.

```
[default]
output = json
region = us-west-1
```

```
aws_access_key_id = myAccessKeyID
aws_secret_access_key = mySecretAccessKey
emr =
     service_role = MyCustomServiceRoleForEMR
     instance_profile = MyCustomServiceRoleForClusterEC2Instances
```

Configuración de roles de IAM para solicitudes de EMRFS a Amazon S3

> **Note**
>
> La capacidad de asignación de roles de EMRFS que se describe en esta página se mejoró con la introducción de Amazon S3 Access Grants en Amazon EMR 6.15.0. Para obtener una solución de control de acceso escalable para los datos en Amazon S3, recomendamos utilizar S3 Access Grants con Amazon EMR.

Cuando una aplicación que se ejecuta en un clúster hace referencia a los datos con el formato `s3://mydata`, Amazon EMR utiliza EMRFS para realizar la solicitud. Para interactuar con Amazon S3, EMRFS asume las políticas de permisos que se adjuntan a su perfil de EC2 instancia de Amazon. Se utiliza el mismo perfil de EC2 instancia de Amazon independientemente del usuario o grupo que ejecute la aplicación o de la ubicación de los datos en Amazon S3.

Si tiene clústeres con varios usuarios que necesitan diferentes niveles de acceso a los datos en Amazon S3 a través de EMRFS, puede establecer una configuración de seguridad con los roles de IAM para EMRFS. EMRFS puede asumir una función de servicio diferente para EC2 las instancias de clúster en función del usuario o grupo que realiza la solicitud o de la ubicación de los datos en Amazon S3. Cada rol de IAM para EMRFS puede tener diferentes permisos para obtener acceso a los datos en Amazon S3. Para obtener más información sobre la función de servicio de las EC2 instancias de clúster, consulte. Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia)

Las versiones 5.10.0 y posteriores de Amazon EMR admiten el uso de roles de IAM personalizados para EMRFS. Si utiliza una versión anterior o tiene requisitos de autorización que los roles de IAM para EMRFS no son capaces de proporcionar, puede crear un proveedor de credenciales personalizado en su lugar. Para obtener más información, consulte Autorización de acceso a los datos de EMRFS en Amazon S3.

Cuando se usa una configuración de seguridad para especificar roles de IAM para EMRFS, debe definir asignaciones de roles. Cada asignación de roles especifica un rol de IAM que corresponde a los identificadores. Dichos identificadores determinan la base para el acceso a Amazon S3 a través de EMRFS. Los identificadores pueden ser usuarios, grupos o prefijos de Amazon S3 que indican una ubicación de datos. Cuando EMRFS realiza una solicitud a Amazon S3, si la solicitud coincide con la base de acceso, EMRFS hace que las EC2 instancias de clúster asuman la función de IAM correspondiente a la solicitud. Para las instancias de clúster, se aplican los permisos de IAM asociados a esa función en lugar de los permisos de IAM adjuntos a la función de servicio. EC2

Los usuarios y grupos de un mapeo de roles son usuarios y grupos de Hadoop que se definen en el clúster. Los usuarios y grupos se transfieren a EMRFS en el contexto de la aplicación que lo utiliza (por ejemplo, una suplantación de usuario de YARN). El prefijo de Amazon S3 puede ser un especificador de bucket de cualquier profundidad (por ejemplo, `s3://amzn-s3-demo-bucket` o `s3://amzn-s3-demo-bucket/myproject/mydata`). Puede especificar varios identificadores dentro de un único mapeo de roles, pero todos deben ser del mismo tipo.

> **Important**
>
> Los roles de IAM para EMRFS proporcionan aislamiento de la aplicación entre los usuarios de la aplicación. No proporcionan aislamiento de nivel del host entre los usuarios del host. Cualquier usuario con acceso a el clúster puede omitir el aislamiento para asumir cualquiera de los roles.

Cuando una aplicación del clúster realiza una solicitud a Amazon S3 a través de EMRFS, EMRFS evalúa las asignaciones de roles en el orden descendente en el que aparecen en la configuración de seguridad. Si una solicitud realizada a través de EMRFS no coincide con ningún identificador, EMRFS vuelve a utilizar la función de servicio para las instancias de clúster. EC2 Por este motivo, le recomendamos que las políticas asociadas a este rol limiten los permisos a Amazon S3. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia).

Configurar roles de

Antes de definir una configuración de seguridad con roles de IAM para EMRFS, planifique y cree los roles y las políticas de permisos que va a adjuntar a los roles. Para obtener más información, consulta ¿Cómo funcionan las funciones de las instancias? EC2 en la Guía del usuario de IAM. Al crear políticas de permisos, le recomendamos que comience con la política gestionada

adjunta a la función Amazon EMR predeterminada y EC2, a continuación, la modifique según sus requisitos. El nombre del rol predeterminado es `EMR_EC2_DefaultRole` y la política administrada predeterminada para editar es `AmazonElasticMapReduceforEC2Role`. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia).

Actualización de políticas de confianza para asumir los permisos del rol

Cada función que utilice EMRFS debe tener una política de confianza que permita que la función Amazon EMR del clúster la asuma. EC2 Del mismo modo, la función Amazon EMR del clúster EC2 debe tener una política de confianza que permita que las funciones de EMRFS la asuman.

En el ejemplo siguiente, la política de confianza se asocia a los roles para EMRFS. La declaración permite que el rol predeterminado de Amazon EMR asuma EC2 el rol. Por ejemplo, si tiene dos roles de EMRFS ficticios, `EMRFSRole_First` y `EMRFSRole_Second`, esta instrucción de la política se añade a las políticas de confianza para cada uno de ellos.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": "arn:aws:iam::123456789012:role/EMR_EC2_DefaultRole",
      "Sid": "AllowSTSAssumerole"
    }
  ]
}
```

Además, en el ejemplo siguiente, se añade la instrucción de la política de confianza a `EMR_EC2_DefaultRole` para permitir que los dos roles de EMRFS ficticios la asuman.

JSON

```
{
  "Version": "2012-10-17",
```

```
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "sts:AssumeRole"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/EMRFSRole_First",
        "arn:aws:iam::123456789012:role/EMRFSRole_Second"
      ],
      "Sid": "AllowSTSAssumerole"
    }
  ]
}
```

Para actualizar la política de confianza de un rol de IAM

Abra la consola de IAM en https://console.aws.amazon.com/iam/.

1. Elija Roles, escriba el nombre del rol en Search (Buscar) y, a continuación, seleccione su Role name (Nombre de rol).
2. Seleccione Trust relationships (Relaciones de confianza), Edit trust relationship (Editar relaciones de confianza).
3. Agregue una instrucción de confianza de acuerdo con Documento de política siguiendo las directrices anteriores y, a continuación, seleccione Actualizar la política de confianza.

Especificación de un rol como un usuario de claves

Si un rol permite el acceso a una ubicación en Amazon S3 que está cifrada mediante una AWS KMS key, asegúrese de que el rol se especifica como un usuario de claves. Esto proporciona al rol permiso para utilizar la clave de KMS. Para obtener más información, consulte Políticas de claves en AWS KMS en la Guía para desarrolladores de AWS Key Management Service .

Definición de una configuración de seguridad con roles de IAM para EMRFS

> **Important**
>
> Si no se aplica ninguna de las funciones de IAM para EMRFS que especifique, EMRFS recurre a la función de Amazon EMR. EC2 Considere la posibilidad de personalizar este rol para restringir los permisos en Amazon S3 según sea necesario para su aplicación y de

especificar después este rol personalizado en lugar de `EMR_EC2_DefaultRole` cuando cree un clúster. Para obtener más información, consulte Personalización de roles de IAM con Amazon EMR y Cómo especificar roles de IAM personalizados al crear un clúster.

Para especificar roles de IAM para solicitudes de EMRFS a Amazon S3 con la consola

1. Cree una configuración de seguridad que especifique los mapeos de roles:
   a. En la consola de Amazon EMR, seleccione Configuraciones de seguridad, Crear.
   b. Escriba un nombre para la configuración de seguridad en el campo Name (Nombre). Este nombre se utiliza para especificar la configuración de seguridad cuando crea un clúster.
   c. Elija Usar roles de IAM para solicitudes de EMRFS a Amazon S3.
   d. Seleccione el rol de IAM que desea aplicar, elija en la lista Base para el acceso un identificador de tipo (Usuarios, Grupos o Prefijos de S3) e ingrese los identificadores correspondientes. Si utiliza varios identificadores, sepárelos con una coma y sin espacios. Para obtener más información acerca de cada tipo de identificador, consulte JSON configuration reference a continuación.
   e. Elija Add role (Añadir rol) para configurar mapeos de roles adicionales, tal y como se describe en el paso anterior.
   f. Defina otras opciones de configuración de seguridad según corresponda y elija Create (Crear). Para obtener más información, consulte Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI.
2. Especifique la configuración de seguridad que creó anteriormente cuando cree un clúster. Para obtener más información, consulte Especificar una configuración de seguridad para un clúster de Amazon EMR.

Para especificar las funciones de IAM para las solicitudes de EMRFS a Amazon S3 mediante el AWS CLI

1. Utilice el comando `aws emr create-security-configuration`, especificando un nombre para la configuración de seguridad y los detalles de configuración de seguridad en formato JSON.

   El comando de ejemplo que se muestra a continuación crea una configuración de seguridad con el nombre `EMRFS_Roles_Security_Configuration`. Se basa en una estructura JSON del

archivo `MyEmrfsSecConfig.json`, que se guarda en el mismo directorio en el que se ejecuta el comando.

```
aws emr create-security-configuration --name EMRFS_Roles_Security_Configuration --
security-configuration file://MyEmrFsSecConfig.json.
```

Utilice las siguientes directrices para la estructura del archivo `MyEmrFsSecConfig.json`. Puede especificar esta estructura junto con estructuras de otras opciones de configuración de seguridad. Para obtener más información, consulte Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI.

A continuación, se muestra un ejemplo de fragmento de JSON para especificar roles de IAM personalizados para EMRFS dentro de una configuración de seguridad. Muestra las asignaciones de roles para los tres tipos de identificadores diferentes, seguidas de una referencia de parámetros.

```
{
  "AuthorizationConfiguration": {
    "EmrFsConfiguration": {
      "RoleMappings": [{
        "Role": "arn:aws:iam::123456789101:role/allow_EMRFS_access_for_user1",
        "IdentifierType": "User",
        "Identifiers": [ "user1" ]
      },{
        "Role": "arn:aws:iam::123456789101:role/
allow_EMRFS_access_to_demo_s3_buckets",
        "IdentifierType": "Prefix",
        "Identifiers": [ "s3://amzn-s3-demo-bucket1/","s3://amzn-s3-demo-
bucket2/" ]
      },{
        "Role": "arn:aws:iam::123456789101:role/allow_EMRFS_access_for_AdminGroup",
        "IdentifierType": "Group",
        "Identifiers": [ "AdminGroup" ]
      }]
    }
  }
}
```

Parámetro	Descripción
`"AuthorizationConfiguration":`	Obligatorio.
`"EmrFsConfiguration":`	Obligatorio. Contiene asignaciones de roles.
`"RoleMappings":`	Obligatorio. Contiene una o más definiciones de asignación de roles. Las asignaciones de roles se evalúan de forma descendente por orden de aparición. Si una asignación de roles se evalúa como true en una invocació n de datos de EMRFS en Amazon S3, no se evalúan más asignaciones de roles y EMRFS utiliza el rol de IAM especificado para la solicitud. Las asignaciones de roles tienen los siguientes parámetros obligatorios:
`"Role":`	Especifica el identificador de ARN de un rol de IAM en el formato `arn:aws:iam::` *`account-id`* `:role/`*`role-name`* . Este es el rol de IAM que asume Amazon EMR si la solicitud de EMRFS a Amazon S3 coincide con alguno de los `Identifiers` especificados.

Parámetro	Descripción
`"IdentifierType":`	Puede ser uno de los siguientes: • `"User"` especifica que los identificadores son uno o más usuarios de Hadoop, que pueden ser usuarios de cuentas de Linux o entidades principales de Kerberos. Cuando la solicitud de EMRFS se origina en el usuario o los usuarios especificados, se asume el rol de IAM. • `"Prefix"` especifica que el identificador es una ubicación de Amazon S3. El rol de IAM se asume para las llamadas a la ubicación o ubicaciones con los prefijos especificados. Por ejemplo, el prefijo `s3://amzn-s3-demo-bucket/` coincide con `s3://amzn-s3-demo-bucket/mydir` y `s3://amzn-s3-demo-bucket/yetanotherdir` . • `"Group"` especifica que los identificadores son uno o más grupos de Hadoop. El rol de IAM se asume si la solicitud proviene de un usuario del grupo o grupos especific ados.
`"Identifiers":`	Especifica uno o más identificadores del tipo de identificador adecuado. Separe varios identificadores con comas sin espacios.

2. Utilice el comando `aws emr create-cluster` para crear un clúster y especificar la configuración de seguridad que creó en el paso anterior.

   En el siguiente ejemplo, se crea un clúster con las aplicaciones de Hadoop básicas predeterminadas instaladas. El clúster usa la configuración de seguridad creada anteriormente `EMRFS_Roles_Security_Configuration` y también usa un rol Amazon EMR personalizado

para EC2`EC2_Role_EMR_Restrict_S3`, que se especifica mediante el `InstanceProfile` argumento del `--ec2-attributes` parámetro.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name MyEmrFsS3RolesCluster \
--release-label emr-7.10.0 --ec2-attributes
 InstanceProfile=EC2_Role_EMR_Restrict_S3,KeyName=MyKey \
--instance-type m5.xlarge --instance-count 3 \
--security-configuration EMRFS_Roles_Security_Configuration
```

Uso de políticas basadas en recursos para el acceso de Amazon EMR a Catálogo de datos de AWS Glue

Si usa AWS Glue junto con Hive, Spark o Presto en Amazon EMR, AWS Glue admite políticas basadas en recursos para controlar el acceso a los recursos del catálogo de datos. Estos recursos incluyen bases de datos, tablas, conexiones y funciones definidas por el usuario. Para obtener más información, consulte Políticas de recursos de AWS Glue en la Guía para desarrolladores de AWS Glue.

Al utilizar políticas basadas en recursos para limitar el acceso a AWS Glue desde Amazon EMR, el principal que especifique en la política de permisos debe ser el ARN de rol asociado al perfil de EC2 instancia que se especifica al crear un clúster. Por ejemplo, para una política basada en recursos adjunta a un catálogo, puede especificar el ARN del rol para el rol de servicio predeterminado para las EC2 instancias de clúster, *`EMR_EC2_DefaultRole`* como el`Principal`, con el formato que se muestra en el siguiente ejemplo:

```
arn:aws:iam::acct-id:role/EMR_EC2_DefaultRole
```

*`acct-id`*Puede ser diferente del ID de la cuenta de AWS Glue. Esto permite el acceso desde clústeres de EMR en diferentes cuentas. Puede especificar varias entidades principales, cada una de ellas desde una cuenta diferente.

Uso de roles de IAM con las aplicaciones que llaman directamente a los servicios de AWS

Las aplicaciones que se ejecutan en las EC2 instancias de un clúster pueden usar el perfil de EC2 instancia para obtener credenciales de seguridad temporales al llamar a AWS los servicios.

Las versiones de Hadoop disponibles con la versión 2.3.0 y posteriores de Amazon EMR ya se han actualizado para utilizar roles de IAM. Si su aplicación se ejecuta estrictamente sobre la arquitectura Hadoop y no llama directamente a ningún servicio AWS, debería funcionar con las funciones de IAM sin modificaciones.

Si su aplicación llama AWS directamente a los servicios, debe actualizarla para aprovechar las funciones de IAM. Esto significa que, en lugar de obtener las credenciales de `/etc/hadoop/conf/core-site.xml` las cuentas de las EC2 instancias del clúster, la aplicación utiliza un SDK para acceder a los recursos mediante las funciones de IAM o invoca los metadatos de la EC2 instancia para obtener las credenciales temporales.

Para acceder a AWS los recursos con funciones de IAM mediante un SDK

- En los siguientes temas se muestra cómo usar varios de ellos para acceder AWS SDKs a las credenciales temporales mediante los roles de IAM. Cada tema comienza con una versión de una aplicación que no utiliza roles de IAM y, a continuación, realiza un recorrido por el proceso de conversión de dicha aplicación para utilizar roles de IAM.
  - Uso de funciones de IAM para EC2 instancias de Amazon con el SDK para Java en la Guía para AWS SDK para Java desarrolladores
  - Uso de roles de IAM para EC2 instancias de Amazon con el SDK para .NET en la Guía para AWS SDK para .NET desarrolladores
  - Uso de funciones de IAM para EC2 instancias de Amazon con el SDK for PHP en la Guía para AWS SDK para PHP desarrolladores
  - Uso de roles de IAM para EC2 instancias de Amazon con el SDK for Ruby en la Guía para AWS SDK para Ruby desarrolladores

Para obtener credenciales temporales a partir de los metadatos de la EC2 instancia

- Llama a la siguiente URL desde una EC2 instancia que se esté ejecutando con la función de IAM especificada, que devolverá las credenciales de seguridad temporales asociadas (AccessKeyId, SecretAccessKey SessionToken, y caducidad). En el ejemplo que

aparece a continuación se utiliza el perfil de instancia predeterminado para Amazon EMR, `EMR_EC2_DefaultRole`.

```
GET http://169.254.169.254/latest/meta-data/iam/security-
credentials/EMR_EC2_DefaultRole
```

Para obtener más información sobre cómo escribir aplicaciones que utilizan funciones de IAM, consulte Conceder acceso a los AWS recursos a las aplicaciones que se ejecutan en EC2 instancias de Amazon.

Para obtener más información acerca de las credenciales de seguridad temporales, consulte Uso de credenciales de seguridad temporales en la guía Uso de credenciales de seguridad temporales.

Cómo permitir a los usuarios y grupos crear y modificar roles

Las entidades principales de IAM (usuarios y grupos) que crean, modifican y especifican roles para un clúster, incluidos los roles predeterminados, deben tener permisos para realizar las acciones siguientes. Para obtener información sobre cada una de las acciones, consulte Acciones en la Referencia de la API de IAM.

- `iam:CreateRole`
- `iam:PutRolePolicy`
- `iam:CreateInstanceProfile`
- `iam:AddRoleToInstanceProfile`
- `iam:ListRoles`
- `iam:GetPolicy`
- `iam:GetInstanceProfile`
- `iam:GetPolicyVersion`
- `iam:AttachRolePolicy`
- `iam:PassRole`

El permiso `iam:PassRole` permite la creación del clúster. Los permisos restantes permiten la creación de las funciones predeterminadas.

Para obtener más información acerca de la actualización de permisos en IAM, consulte Cambio de los permisos de un usuario en la Guía del usuario de IAM.

Ejemplos de políticas de Amazon EMR basadas en identidades

De forma predeterminada, los usuarios y roles no tienen permiso para crear ni modificar los recursos de Amazon EMR. Tampoco pueden realizar tareas con la AWS API AWS Management Console AWS CLI, o. Un administrador de IAM debe crear políticas de IAM que concedan permisos a los usuarios y a los roles para realizar operaciones de la API concretas en los recursos especificados que necesiten. El administrador debe adjuntar esas políticas a los usuarios o grupos que necesiten esos permisos.

Para obtener más información acerca de cómo crear una política basada en identidad de IAM con estos documentos de políticas de JSON de ejemplo, consulte Creación de políticas en la pestaña JSON en la Guía del usuario de IAM.

Temas

- Prácticas recomendadas en materia de políticas para Amazon EMR
- Cómo permitir a los usuarios consultar sus propios permisos
- Políticas administradas por Amazon EMR
- Políticas de IAM para el acceso basado en etiquetas para clústeres y cuadernos de EMR
- Denegar la ModifyInstanceGroup acción en Amazon EMR
- Solución de problemas de identidad y acceso de Amazon EMR

Prácticas recomendadas en materia de políticas para Amazon EMR

Las políticas basadas en identidad son muy eficaces. Determinan si alguien puede crear o eliminar los recursos de Amazon EMR de su cuenta o acceder a ellos. Estas acciones pueden conllevar costes para tu AWS cuenta. Siga estas directrices y recomendaciones al crear o editar políticas basadas en identidades:

- Comience a utilizar las políticas AWS gestionadas: para empezar a utilizar Amazon EMR rápidamente, utilice las políticas AWS gestionadas para conceder a sus empleados los permisos que necesitan. Estas políticas ya están disponibles en su cuenta, y las mantiene y actualiza AWS. Para obtener más información, consulte Cómo empezar a usar permisos con políticas AWS administradas en la Guía del usuario de IAM y. Políticas administradas por Amazon EMR

- Conceder privilegios mínimos: al crear políticas personalizadas, conceda solo los permisos necesarios para llevar a cabo una tarea. Comience con un conjunto mínimo de permisos y conceda permisos adicionales según sea necesario. Por lo general, es más seguro que comenzar con permisos demasiado tolerantes e intentar hacerlos más estrictos más adelante. Para obtener más información, consulte Conceder privilegios mínimos en la Guía del usuario de IAM.
- Habilitar MFA para operaciones confidenciales: para mayor seguridad, obligue a los usuarios de que utilicen la autenticación multifactor (MFA) para acceder a recursos u operaciones de API confidenciales. Para obtener más información, consulte Uso de la autenticación multifactor (MFA) en AWS en la Guía del usuario de IAM.
- Utilizar condiciones de política para mayor seguridad: en la medida en que sea práctico, defina las condiciones en las que sus políticas basadas en identidad permitan el acceso a un recurso. Por ejemplo, puede escribir condiciones para especificar un rango de direcciones IP permitidas desde el que debe proceder una solicitud. También puede escribir condiciones para permitir solicitudes solo en un intervalo de hora o fecha especificado o para solicitar el uso de SSL o MFA. Para obtener más información, consulte Elementos de la política JSON de IAM: condición en la Guía del usuario de IAM.

Cómo permitir a los usuarios consultar sus propios permisos

En este ejemplo, se muestra cómo podría crear una política que permita a los usuarios de IAM ver las políticas administradas e insertadas que se adjuntan a la identidad de sus usuarios. Esta política incluye permisos para completar esta acción en la consola o mediante programación mediante la AWS CLI API o. AWS

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ViewOwnUserInfo",
      "Effect": "Allow",
      "Action": [
        "iam:GetUser",
        "iam:GetUserPolicy",
        "iam:ListAttachedUserPolicies",
        "iam:ListGroupsForUser",
        "iam:ListUserPolicies"
```

```
      ],
      "Resource": [
        "arn:aws:iam::*:user/${aws:username}"
      ]
    },
    {
      "Sid": "NavigateInConsole",
      "Effect": "Allow",
      "Action": [
        "iam:GetGroupPolicy",
        "iam:GetPolicy",
        "iam:GetPolicyVersion",
        "iam:ListAttachedGroupPolicies",
        "iam:ListGroupPolicies",
        "iam:ListGroups",
        "iam:ListPolicies",
        "iam:ListPolicyVersions",
        "iam:ListUsers"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Políticas administradas por Amazon EMR

La forma más sencilla de conceder acceso completo o acceso de solo lectura a las acciones de Amazon EMR requeridas consiste en utilizar las políticas administradas de IAM para Amazon EMR. Las políticas administradas ofrecen el beneficio de que se actualizan automáticamente si cambian los requisitos de permisos. Si utiliza políticas insertadas, pueden producirse cambios en los servicios que provoquen la aparición de errores de permisos.

Amazon EMR dejará de utilizar las políticas administradas existentes (políticas de la versión 1) en favor de las nuevas políticas administradas (políticas de la versión 2). Se ha reducido el alcance de las nuevas políticas gestionadas para alinearlas con las mejores prácticas. AWS Una vez que las políticas administradas de la versión 1 estén obsoletas, no podrá adjuntarlas a ningún rol o usuario de IAM nuevo. Los roles y usuarios existentes que usan políticas obsoletas pueden seguir usándolas. Las políticas administradas de la versión 2 restringen el acceso mediante etiquetas. Solo permiten acciones específicas de Amazon EMR y requieren recursos de clúster etiquetados con una clave

específica de EMR. Le recomendamos que revise detenidamente la documentación antes de utilizar las nuevas políticas de la versión 2.

Las políticas de la versión 1 se marcarán como obsoletas con un icono de advertencia a su lado en la lista Políticas de la consola de IAM. Las políticas obsoletas tienen las siguientes características:

- Siguen funcionando para todos los usuarios, grupos y roles asociados en ese momento. Ningún elemento deja de funcionar.
- No pueden asociarse a ningún usuario, grupo o rol nuevo. Si separa una política de una entidad actual, no puede volver a asociarla.
- Después de separar una política de la versión 1 de todas las entidades actuales, la política dejará de estar visible y ya no podrá utilizarse.

La siguiente tabla resume los cambios entre las políticas actuales (versión 1) y las políticas de la versión 2.

Cambios en las políticas administradas por Amazon EMR

Tipo de política	Nombres de las políticas	Propósito de la política	Cambios en la política de la versión 2
Rol de servicio de EMR predeterminado y política administrada adjunta	Nombre del rol: EMR_DefaultRole Política V1 (quedará obsoleta): AmazonElasticMapReduceRole(Rol de servicio EMR) Nombre de la política de la versión 2 (con ámbito de aplicación reducido): **AmazonEMRServicePolicy_v2**	Permite a Amazon EMR llamar a otros AWS servicios en su nombre al aprovisionar recursos y realizar acciones a nivel de servicio. Este rol es necesario para todos los clústeres.	La política añade el nuevo permiso `"ec2:DescribeInstanceTypeOfferings"` . Esta operación de API devuelve una lista de tipos de instancias compatibles con una lista de zonas de disponibilidad determinadas.

Tipo de política	Nombres de las políticas	Propósito de la política	Cambios en la política de la versión 2
Política administrada por IAM para un acceso integral a Amazon EMR por parte del usuario, rol o grupo asociado	Nombre de la política de la versión 2 (con ámbito de aplicación): `AmazonEMRServicePolicy_v2`	Concede a los usuarios permisos completos para realizar acciones de EMR. Incluye iam: permisos para recursos. PassRole	La política agrega un requisito previo según el cual los usuarios deben agregar etiquetas de usuario a los recursos antes de poder usar esta política. Consulte Etiquetado de recursos para usar políticas administradas. iam: la PassRole acción requiere la PassedToService condición iam: establecida en el servicio especificado. El acceso a Amazon EC2, Amazon S3 y otros servicios no está permitido de forma predeterminada. Consulte Política administrada por IAM para un acceso total (política predeterminada administrada de la versión 2).

Tipo de política	Nombres de las políticas	Propósito de la política	Cambios en la política de la versión 2
Política administrada por IAM para el acceso de solo lectura por parte del usuario, rol o grupo asociado	Política de la versión 1 (quedará obsoleta): `AmazonElasticMapReduceReadOnlyAccess` Nombre de la política de la versión 2 (con ámbito de aplicación): `AmazonEMRReadOnlyAccessPolicy_v2`	Concede a los usuarios permisos de solo lectura para realizar acciones de Amazon EMR.	Los permisos permiten únicamente acciones específicas de solo lectura de elasticmapreduce. El acceso a Amazon S3 no está permitido de forma predeterminada. Consulte Política administrada por IAM para un acceso de solo lectura (política predeterminada administrada de la versión 2).

Tipo de política	Nombres de las políticas	Propósito de la política	Cambios en la política de la versión 2
Rol de servicio de EMR predeterminado y política administrada adjunta	Nombre del rol: EMR_DefaultRole Política V1 (quedará obsoleta): AmazonElasticMapReduceRole(Rol de servicio EMR) Nombre de la política de la versión 2 (con ámbito de aplicación reducido): `AmazonEMRServicePolicy_v2`	Permite a Amazon EMR llamar a otros AWS servicios en su nombre al aprovisionar recursos y realizar acciones a nivel de servicio. Este rol es necesario para todos los clústeres.	El rol de servicio de la versión 2 y la política predeterminada de la versión 2 sustituyen a la política y al rol obsoletos. La política agrega un requisito previo según el cual los usuarios deben agregar etiquetas de usuario a los recursos antes de poder usar esta política. Consulte Etiquetado de recursos para usar políticas administradas. Consulte Rol de servicio para Amazon EMR (rol de EMR).

Tipo de política	Nombres de las políticas	Propósito de la política	Cambios en la política de la versión 2
Función de servicio para EC2 instancias de clúster (perfil de instancia) EC2	Nombre del rol: EMR_ _ EC2 DefaultRole Nombre de la política obsoleta: Role AmazonElasticMapReducefor EC2	Permite que las aplicaciones que se ejecutan en un clúster de EMR accedan a otros recursos de AWS , como Amazon S3. Por ejemplo, si ejecuta trabajos de Apache Spark que procesan datos de Amazon S3, la política debe permitir el acceso a dichos recursos.	Tanto el rol predeterminado como la política predeterminada están en vías de quedar obsoletos. No hay ninguna política o función gestionada a AWS predeterminada que la sustituya . Debe proporcionar una política basada en los recursos o en la identidad. Esto significa que, de forma predeterminada, las aplicaciones que se ejecuten en un clúster de EMR no tienen acceso a Amazon S3 ni a ningún otro recurso, a menos que las agregue manualmente a la política. Consulte Política administrada y rol predeterminados.

Tipo de política	Nombres de las políticas	Propósito de la política	Cambios en la política de la versión 2
Otras políticas EC2 de roles de servicio	Nombres de políticas actuales: AmazonElasticMapReduceforAutoScalingRole AmazonElasticMapReduceEditorsRole, Amazon EMRCleanup Policy	Proporciona los permisos que Amazon EMR necesita para acceder a otros AWS recursos y realizar acciones si utiliza el escalado automático, los cuadernos o para limpiar recursos. EC2	No hay cambios para la versión 2.

Proteger iam: PassRole

Las políticas administradas de forma predeterminada con todos los permisos de Amazon EMR incorporan configuraciones de seguridad `iam:PassRole`, entre las que se incluyen las siguientes:

- Permisos `iam:PassRole` solo para roles específicos de Amazon EMR predeterminados.
- `iam:PassedToServicecondiciones` que le permiten usar la política solo con AWS servicios específicos, como `elasticmapreduce.amazonaws.com` `yec2.amazonaws.com`.

Puede ver la versión JSON de las políticas Amazon EMRFull AccessPolicy _v2 y Amazon EMRService Policy_v2 en la consola de IAM. Le recomendamos crear nuevos clústeres con las políticas administradas de la versión 2.

Para crear políticas personalizadas, le recomendamos que comience a partir de las políticas administradas y las edite de acuerdo con sus requisitos.

Para obtener información sobre cómo asociar políticas a los usuarios (entidades principales), consulte Uso de políticas administradas con la AWS Management Console en la Guía del usuario de IAM.

Etiquetado de recursos para usar políticas administradas

Amazon EMRService Policy_v2 y Amazon EMRFull AccessPolicy _v2 dependen del acceso limitado a los recursos que Amazon EMR aprovisiona o utiliza. La reducción del alcance se logra restringiendo el acceso únicamente a los recursos que tienen una etiqueta de usuario predefinida

asociada a ellos. Si utiliza cualquiera de estas dos políticas, debe pasar la etiqueta de usuario predefinida `for-use-with-amazon-emr-managed-policies = true` al aprovisionar el clúster. A continuación, Amazon EMR propagará automáticamente esa etiqueta. Además, debe agregar una etiqueta de usuario a los recursos que se enumeran en la siguiente sección. Si utiliza la consola de Amazon EMR para lanzar el clúster, consulte Consideraciones sobre el uso de la consola de Amazon EMR para lanzar clústeres con políticas administradas de la versión 2.

Para usar políticas administradas, pase la etiqueta de usuario `for-use-with-amazon-emr-managed-policies = true` al aprovisionar un clúster con la CLI, el SDK u otro método.

Al pasar la etiqueta, Amazon EMR la propaga a los volúmenes de ENI, EC2 instancia y EBS de la subred privada que crea. Amazon EMR también etiqueta automáticamente los grupos de seguridad que cree. Sin embargo, si desea que Amazon EMR se lance con un grupo de seguridad determinado, debe etiquetarlo. En el caso de los recursos que no haya creado Amazon EMR, debe agregar etiquetas a esos recursos. Por ejemplo, debe etiquetar EC2 las subredes de Amazon, los grupos de EC2 seguridad (si no los creó Amazon EMR) VPCs y (si desea que Amazon EMR cree grupos de seguridad). Para lanzar clústeres con políticas gestionadas en la versión 2 VPCs, debe etiquetarlos VPCs con la etiqueta de usuario predefinida. Consulte, Consideraciones sobre el uso de la consola de Amazon EMR para lanzar clústeres con políticas administradas de la versión 2.

Etiquetado propagado especificado por el usuario

Amazon EMR etiqueta los recursos que crea mediante las etiquetas de Amazon EMR que especifique al crear un clúster. Amazon EMR aplica etiquetas a los recursos que cree durante la vida útil del clúster.

Amazon EMR propaga las etiquetas de usuario de los siguientes recursos:

- ENI de Subred privada (interfaces de red elásticas de acceso a servicios)
- EC2 Instancias
- Volúmenes de EBS
- EC2 Plantilla de lanzamiento

Grupos de seguridad etiquetados automáticamente

Amazon EMR etiqueta los grupos de EC2 seguridad que crea con la etiqueta necesaria para las políticas gestionadas de Amazon EMR en la versión 2`for-use-with-amazon-emr-managed-policies`, independientemente de las etiquetas que especifique en el comando create cluster. En el caso de un grupo de seguridad que se creó antes de la introducción de las políticas administradas

de la versión 2, Amazon EMR no etiqueta automáticamente el grupo de seguridad. Si desea utilizar políticas administradas de la versión 2 con los grupos de seguridad predeterminados que ya existen en la cuenta, debe etiquetar los grupos de seguridad manualmente con `for-use-with-amazon-emr-managed-policies = true`.

Recursos de clúster etiquetados manualmente

Debe etiquetar manualmente algunos recursos del clúster para que los roles predeterminados de Amazon EMR puedan acceder a ellos.

- Debe etiquetar manualmente los grupos EC2 de seguridad y EC2 las subredes con la etiqueta de política administrada de Amazon EMR. `for-use-with-amazon-emr-managed-policies`
- Debe etiquetar manualmente una VPC si quiere que Amazon EMR cree grupos de seguridad predeterminados. EMR intentará crear un grupo de seguridad con la etiqueta específica si el grupo de seguridad predeterminado aún no existe.

Amazon EMR etiqueta automáticamente los siguientes recursos:

- Grupos de seguridad creados por EMR EC2

Debe etiquetar de forma manual los siguientes recursos:

- EC2 Subred
- EC2 Grupos de seguridad

De forma opcional, puede etiquetar de forma manual los siguientes recursos:

- VPC: solo cuando desee que Amazon EMR cree grupos de seguridad

Consideraciones sobre el uso de la consola de Amazon EMR para lanzar clústeres con políticas administradas de la versión 2

Puede aprovisionar clústeres con políticas administradas de la versión 2 mediante la consola de Amazon EMR. Estas son algunas consideraciones que debe tener en cuenta al utilizar la consola para lanzar clústeres de Amazon EMR.

- No es necesario pasar la etiqueta predefinida. Amazon EMR agrega automáticamente la etiqueta y la propaga a los componentes correspondientes.

- En el caso de los componentes que deben etiquetarse manualmente, la antigua consola de Amazon EMR intenta etiquetarlos automáticamente si tiene los permisos necesarios para etiquetar los recursos. Si no tiene los permisos para etiquetar los recursos o si desea utilizar la consola, pida al administrador que etiquete esos recursos.
- No puede lanzar clústeres con políticas administradas de la versión 2, a menos que se cumplan todos los requisitos previos.
- La consola antigua de Amazon EMR le muestra qué recursos (VPC/subredes) deben etiquetarse.

Política administrada por IAM para un acceso total (política predeterminada administrada de la versión 2) para Amazon EMR

Las políticas administradas por defecto de EMR con el ámbito de aplicación de la versión 2 otorgan privilegios de acceso específicos a los usuarios. Requieren una etiqueta de recurso de Amazon EMR predefinida y claves de condición de `iam:PassRole` para los recursos que utiliza Amazon EMR, como la `Subnet` y el `SecurityGroup` que utiliza para lanzar el clúster.

Para conceder todas las acciones necesarias dentro del ámbito de aplicación de Amazon EMR, adjunte la política administrada `AmazonEMRFullAccessPolicy_v2`. Esta política administrada predeterminada actualizada sustituye a la política administrada `AmazonElasticMapReduceFullAccess`.

`AmazonEMRFullAccessPolicy_v2` depende del acceso limitado a los recursos que Amazon EMR aprovisiona o utiliza. Cuando utilice esta política, tendrá que pasar la etiqueta de usuario `for-use-with-amazon-emr-managed-policies = true` al aprovisionar el clúster. Amazon EMR propagará automáticamente la etiqueta. Además, es posible que tenga que añadir manualmente una etiqueta de usuario a tipos específicos de recursos, como grupos de EC2 seguridad que Amazon EMR no creó. Para obtener más información, consulte Etiquetado de recursos para usar políticas administradas.

La política `AmazonEMRFullAccessPolicy_v2` protege los recursos de la siguiente manera:

- Requiere que los recursos se etiqueten con la etiqueta de políticas administradas de Amazon EMR predefinida `for-use-with-amazon-emr-managed-policies` para la creación de clústeres y el acceso a Amazon EMR.
- Restringe la acción `iam:PassRole` a roles predeterminados específicos y el acceso `iam:PassedToService` a servicios específicos.
- Ya no proporciona acceso a Amazon EC2, Amazon S3 y otros servicios de forma predeterminada.

El contenido de esta política se muestra a continuación.

> **Note**
>
> También puede utilizar el enlace de la consola AmazonEMRFullAccessPolicy_v2 para ver esta política.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "RunJobFlowExplicitlyWithEMRManagedTag",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:RunJobFlow"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/for-use-with-amazon-emr-managed-policies": "true"
        }
      }
    },
    {
      "Sid": "ElasticMapReduceActions",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:AddInstanceFleet",
        "elasticmapreduce:AddInstanceGroups",
        "elasticmapreduce:AddJobFlowSteps",
        "elasticmapreduce:AddTags",
        "elasticmapreduce:CancelSteps",
        "elasticmapreduce:CreateEditor",
        "elasticmapreduce:CreatePersistentAppUI",
        "elasticmapreduce:CreateSecurityConfiguration",
        "elasticmapreduce:DeleteEditor",
        "elasticmapreduce:DeleteSecurityConfiguration",
```

```
        "elasticmapreduce:DescribeCluster",
        "elasticmapreduce:DescribeEditor",
        "elasticmapreduce:DescribeJobFlows",
        "elasticmapreduce:DescribePersistentAppUI",
        "elasticmapreduce:DescribeSecurityConfiguration",
        "elasticmapreduce:DescribeStep",
        "elasticmapreduce:DescribeReleaseLabel",
        "elasticmapreduce:GetBlockPublicAccessConfiguration",
        "elasticmapreduce:GetManagedScalingPolicy",
        "elasticmapreduce:GetPersistentAppUIPresignedURL",
        "elasticmapreduce:GetAutoTerminationPolicy",
        "elasticmapreduce:ListBootstrapActions",
        "elasticmapreduce:ListClusters",
        "elasticmapreduce:ListEditors",
        "elasticmapreduce:ListInstanceFleets",
        "elasticmapreduce:ListInstanceGroups",
        "elasticmapreduce:ListInstances",
        "elasticmapreduce:ListSecurityConfigurations",
        "elasticmapreduce:ListSteps",
        "elasticmapreduce:ListSupportedInstanceTypes",
        "elasticmapreduce:ModifyCluster",
        "elasticmapreduce:ModifyInstanceFleet",
        "elasticmapreduce:ModifyInstanceGroups",
        "elasticmapreduce:OpenEditorInConsole",
        "elasticmapreduce:PutAutoScalingPolicy",
        "elasticmapreduce:PutBlockPublicAccessConfiguration",
        "elasticmapreduce:PutManagedScalingPolicy",
        "elasticmapreduce:RemoveAutoScalingPolicy",
        "elasticmapreduce:RemoveManagedScalingPolicy",
        "elasticmapreduce:RemoveTags",
        "elasticmapreduce:SetTerminationProtection",
        "elasticmapreduce:StartEditor",
        "elasticmapreduce:StopEditor",
        "elasticmapreduce:TerminateJobFlows",
        "elasticmapreduce:ViewEventsFromAllClustersInConsole"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "ViewMetricsInEMRConsole",
      "Effect": "Allow",
      "Action": [
```

```
        "cloudwatch:GetMetricStatistics"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "PassRoleForElasticMapReduce",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/EMR_DefaultRole",
        "arn:aws:iam::*:role/EMR_DefaultRole_V2"
      ],
      "Condition": {
        "StringLike": {
          "iam:PassedToService": "elasticmapreduce.amazonaws.com*"
        }
      }
    },
    {
      "Sid": "PassRoleForEC2",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/EMR_EC2_DefaultRole"
      ],
      "Condition": {
        "StringLike": {
          "iam:PassedToService": "ec2.amazonaws.com*"
        }
      }
    },
    {
      "Sid": "PassRoleForAutoScaling",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
```

```
        "arn:aws:iam::*:role/EMR_AutoScaling_DefaultRole"
      ],
      "Condition": {
        "StringLike": {
          "iam:PassedToService": "application-autoscaling.amazonaws.com*"
        }
      }
    },
    {
      "Sid": "ElasticMapReduceServiceLinkedRole",
      "Effect": "Allow",
      "Action": [
        "iam:CreateServiceLinkedRole"
      ],
      "Resource": [
        "arn:aws:iam::*:role/aws-service-role/elasticmapreduce.amazonaws.com*/
AWSServiceRoleForEMRCleanup*"
      ],
      "Condition": {
        "StringEquals": {
          "iam:AWSServiceName": [
            "elasticmapreduce.amazonaws.com",
            "elasticmapreduce.amazonaws.com.cn"
          ]
        }
      }
    },
    {
      "Sid": "ConsoleUIActions",
      "Effect": "Allow",
      "Action": [
        "ec2:DescribeAccountAttributes",
        "ec2:DescribeAvailabilityZones",
        "ec2:DescribeImages",
        "ec2:DescribeKeyPairs",
        "ec2:DescribeNatGateways",
        "ec2:DescribeRouteTables",
        "ec2:DescribeSecurityGroups",
        "ec2:DescribeSubnets",
        "ec2:DescribeVpcs",
        "ec2:DescribeVpcEndpoints",
        "s3:ListAllMyBuckets",
        "iam:ListRoles"
      ],
```

```
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Política administrada por IAM para obtener acceso completo (en vías de quedar obsoleta)

Las políticas gestionadas `AmazonElasticMapReduceFullAccess` y `AmazonEMRFullAccessPolicy_v2` AWS Identity and Access Management (IAM) otorgan todas las acciones necesarias para Amazon EMR y otros servicios.

> **Important**
>
> La política administrada `AmazonElasticMapReduceFullAccess` está en vías de quedar obsoleta y ya no se recomienda su uso con Amazon EMR. En su lugar, utilice `AmazonEMRFullAccessPolicy_v2`. Cuando el servicio de IAM haga que la política de la versión 1 quede obsoleta, ya no podrá asociarla a ningún rol. Sin embargo, puede adjuntar un rol existente a un clúster incluso si ese rol usa la política obsoleta.

Las políticas administradas de forma predeterminada con todos los permisos de Amazon EMR incorporan configuraciones de seguridad `iam:PassRole`, entre las que se incluyen las siguientes:

- Permisos `iam:PassRole` solo para roles específicos de Amazon EMR predeterminados.
- `iam:PassedToService`condiciones que le permiten usar la política solo con AWS servicios específicos, como `elasticmapreduce.amazonaws.com` y. `ec2.amazonaws.com`

Puede ver la versión JSON de las políticas Amazon EMRFull AccessPolicy _v2 y Amazon EMRService Policy_v2 en la consola de IAM. Le recomendamos crear nuevos clústeres con las políticas administradas de la versión 2.

Puede ver el contenido de la política de la versión 1, que está obsoleta, en at. AWS Management Console `AmazonElasticMapReduceFullAccess` La `ec2:TerminateInstances` acción de la política otorga permiso al usuario o rol para cancelar cualquiera de las EC2 instancias de Amazon asociadas a la cuenta de IAM. Esto incluye instancias que no forman parte de un clúster de EMR.

Política administrada por IAM para el acceso de solo lectura (política predeterminada administrada de la versión 2) para Amazon EMR

Para conceder privilegios de solo lectura a Amazon EMR, adjunte la política gestionada Amazon EMRRead OnlyAccessPolicy _v2. Esta política administrada predeterminada reemplaza a la política administrada `AmazonElasticMapReduceReadOnlyAccess`. El contenido de esta instrucción de la política se muestra en el siguiente fragmento: En comparación con la política `AmazonElasticMapReduceReadOnlyAccess`, la política `AmazonEMRReadOnlyAccessPolicy_v2` no utiliza caracteres comodín para el elemento `elasticmapreduce`. En su lugar, la política de la versión 2 predeterminada determina el ámbito de aplicación de las acciones `elasticmapreduce` permitidas.

> **Note**
>
> También puede utilizar el AWS Management Console enlace para ver la política `AmazonEMRReadOnlyAccessPolicy_v2`.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "ElasticMapReduceActions",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:DescribeCluster",
        "elasticmapreduce:DescribeEditor",
        "elasticmapreduce:DescribeJobFlows",
        "elasticmapreduce:DescribeSecurityConfiguration",
        "elasticmapreduce:DescribeStep",
        "elasticmapreduce:DescribeReleaseLabel",
        "elasticmapreduce:GetBlockPublicAccessConfiguration",
        "elasticmapreduce:GetManagedScalingPolicy",
        "elasticmapreduce:GetAutoTerminationPolicy",
        "elasticmapreduce:ListBootstrapActions",
        "elasticmapreduce:ListClusters",
        "elasticmapreduce:ListEditors",
        "elasticmapreduce:ListInstanceFleets",
```

```
        "elasticmapreduce:ListInstanceGroups",
        "elasticmapreduce:ListInstances",
        "elasticmapreduce:ListSecurityConfigurations",
        "elasticmapreduce:ListSteps",
        "elasticmapreduce:ListSupportedInstanceTypes",
        "elasticmapreduce:ViewEventsFromAllClustersInConsole"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "ViewMetricsInEMRConsole",
      "Effect": "Allow",
      "Action": [
        "cloudwatch:GetMetricStatistics"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

Política administrada por IAM para obtener acceso de solo lecturas (en vías de quedar obsoleta)

La política administrada `AmazonElasticMapReduceReadOnlyAccess` está en vías de quedar obsoleta. No puede adjuntar esta política al lanzar nuevos clústeres. `AmazonElasticMapReduceReadOnlyAccess` se ha sustituido por `AmazonEMRReadOnlyAccessPolicy_v2`, que ahora es la política administrada predeterminada de Amazon EMR. El contenido de esta instrucción de la política se muestra en el siguiente fragmento: Los caracteres comodín para el elemento `elasticmapreduce` especifican que solo se permiten las acciones que empiezan por las cadenas especificadas. Tenga en cuenta que, dado que esta política no deniega acciones explícitamente, se puede seguir utilizando una instrucción de política distinta para otorgar acceso a acciones especificadas.

> **Note**
>
> También puede utilizar el AWS Management Console para ver la política.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:Describe*",
        "elasticmapreduce:List*",
        "elasticmapreduce:ViewEventsFromAllClustersInConsole",
        "s3:GetObject",
        "s3:ListAllMyBuckets",
        "s3:ListBucket",
        "sdb:Select",
        "cloudwatch:GetMetricStatistics"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowELASTICMAPREDUCEDescribe"
    }
  ]
}
```

AWS política gestionada: EMRDescribe ClusterPolicyFor EMRAL

No puede asociar `EMRDescribeClusterPolicyForEMRWAL` a sus entidades IAM. Esta política está adjunta a un rol vinculado a servicios que permite a Amazon EMR realizar acciones en su nombre. Para obtener más información sobre este rol vinculado a servicios, consulte Uso de roles vinculados a servicios con Amazon EMR para el registro de escritura anticipada.

Esta política concede permisos de solo lectura que permiten al servicio WAL de Amazon EMR encontrar y devolver el estado de un clúster. Para obtener más información sobre Amazon EMR WAL, consulte Registros de escritura anticipada (WAL) para Amazon EMR.

Detalles de los permisos

Esta política incluye los permisos siguientes:

- `emr` – Permite a las entidades principales describir el estado de clúster de Amazon EMR. Esto es necesario para que Amazon EMR pueda confirmar la finalización de un clúster y, después de treinta días, limpiar todos los registros de WAL que haya dejado el clúster.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:DescribeCluster"
      ],
      "Resource": [
        "*"
      ],
      "Sid": "AllowELASTICMAPREDUCEDescribecluster"
    }
  ]
}
```

AWS políticas gestionadas para Amazon EMR

Una política AWS administrada es una política independiente creada y administrada por. AWS AWS Las políticas administradas están diseñadas para proporcionar permisos para muchos casos de uso comunes, de modo que pueda empezar a asignar permisos a usuarios, grupos y funciones.

Ten en cuenta que es posible que las políticas AWS administradas no otorguen permisos con privilegios mínimos para tus casos de uso específicos, ya que están disponibles para que los usen todos los AWS clientes. Se recomienda definir políticas administradas por el cliente específicas para sus casos de uso a fin de reducir aún más los permisos.

No puedes cambiar los permisos definidos en AWS las políticas administradas. Si AWS actualiza los permisos definidos en una política AWS administrada, la actualización afecta a todas las identidades principales (usuarios, grupos y roles) a las que está asociada la política. AWS es más probable que actualice una política AWS administrada cuando Servicio de AWS se lance una nueva o cuando estén disponibles nuevas operaciones de API para los servicios existentes.

Para obtener más información, consulte Políticas administradas de AWS en la Guía del usuario de IAM.

Amazon EMR actualiza las políticas gestionadas AWS

Consulte los detalles sobre las actualizaciones de las políticas AWS gestionadas para Amazon EMR desde que este servicio comenzó a realizar el seguimiento de estos cambios.

Cambio	Descripción	Fecha
AmazonEMRServicePolicy_v2 : actualización de una política actual	Se agregó `ec2:CreateVpcEndpoint` y se `ec2:CreateTags` requiere para una experiencia óptima, a partir de la versión 7.5.0 de Amazon EMR. `ec2:ModifyVpcEndpoint`	4 de marzo de 2025
AmazonEMRServicePolicy_v2 : actualización de una política actual	Se agregaron `elasticmapreduce:CreatePersistentAppUI` `elasticmapreduce:DescribePersistentAppUI` , y. `elasticmapreduce:GetPersistentAppUIPresignedURL`	28 de febrero de 2025
EMRDescribeClusterPolicyForEMRWAL : política nueva	Se agregó una nueva política para que Amazon EMR pueda determinar el estado del clúster para la limpieza de WAL treinta días después de la finalización del clúster.	10 de agosto de 2023
AmazonEMRFullAccessPolicy_v2 y **AmazonEMRReadOnlyAccessPoli**	Se añadieron `elasticmapreduce:DescribeReleaseLabel` y `elasticma`	21 de abril de 2022

Cambio	Descripción	Fecha
cy_v2 : actualización de una política existente	preduce:GetAutoTerminationPolicy .	
AmazonEMRFullAccessPolicy_v2 : actualización de una política actual	Se agregó ec2:DescribeImages para Uso de una AMI personalizada para ofrecer más flexibilidad a la configuración del clúster de Amazon EMR.	15 de febrero de 2022
Políticas administradas por Amazon EMR	Se actualizó para aclarar el uso de etiquetas de usuario predefinidas. Se agregó una sección sobre el uso de la AWS consola para lanzar clústeres con políticas administradas en la versión 2.	29 de septiembre de 2021
AmazonEMRFullAccessPolicy_v2 : actualización de una política actual	Se modificaron las acciones PassRoleForEC2 y PassRoleForAutoScaling para que usen el operador de condición StringLike y que así coincidan con "iam:PassedToService":"application-autoscaling.amazonaws.com*" y "iam:PassedToService":"ec2.amazonaws.com*" , respectivamente.	20 de mayo de 2021

Cambio	Descripción	Fecha
AmazonEMRFullAccessPolicy_v2: actualización de una política actual	Se ha eliminado la acción no válida `s3:ListBuckets` y se ha sustituido por la acción `s3:ListAllMyBuckets`. Se actualizó la creación de roles vinculados a un servicio (SLR) para que se limite explícitamente al único rol que tiene Amazon EMR con principios de servicio explícitos. Los SLRs que se pueden crear son exactamente los mismos que antes de este cambio.	23 de marzo de 2021

Cambio	Descripción	Fecha
AmazonEMRFullAccessPolicy_v2: política nueva	Amazon EMR ha agregado nuevos permisos para limitar el acceso a los recursos y para agregar un requisito previo según el cual los usuarios deben agregar una etiqueta de usuario predefinida a los recursos antes de poder utilizar las políticas administradas de Amazon EMR. La acción `iam:PassRole` requiere que se establezca una condición `iam:PassedToService` en un servicio específico. El acceso a Amazon EC2, Amazon S3 y otros servicios no está permitido de forma predeterminada.	11 de marzo de 2021
AmazonEMRServicePolicy_v2: política nueva	Agrega un requisito previo según el cual los usuarios deben agregar etiquetas de usuario a los recursos antes de poder utilizar esta política.	11 de marzo de 2021
AmazonEMRReadOnlyAccessPolicy_v2: política nueva	Los permisos permiten únicamente acciones específicas de solo lectura de elasticmapreduce. El acceso a Amazon S3 no está permitido de forma predeterminada.	11 de marzo de 2021

Cambio	Descripción	Fecha
Amazon EMR ha comenzado a hacer un seguimiento de los cambios	Amazon EMR comenzó a realizar un seguimiento de los cambios de sus políticas AWS gestionadas.	11 de marzo de 2021

Políticas de IAM para el acceso basado en etiquetas para clústeres y cuadernos de EMR

Puede utilizar condiciones en su política basada en identidades para controlar el acceso a clústeres y blocs de notas de EMR basándose en etiquetas.

Para obtener más información sobre la adición de etiquetas a los clústeres de EMR, consulte Etiquetado de clústeres de Amazon EMR.

Los siguientes ejemplos muestran distintos supuestos y formas de utilizar los operadores de condición con las claves de condición de Amazon EMR. Estas instrucciones de política de IAM tienen fines demostrativos y no deben utilizarse en entornos de producción. Existen varias maneras de combinar las instrucciones de políticas para conceder y denegar permisos de acuerdo con sus requisitos. Para obtener más información sobre la planificación y las pruebas de políticas de IAM, consulte la Guía del usuario de IAM.

> **Important**
>
> La denegación de permisos explícita para acciones de etiquetado de acciones es un factor importante. Esto impide que los usuarios etiqueten un recurso y, de esta forma, se concedan a sí mismos permisos que usted no tenía previsto conceder. Si no deniega las acciones de etiquetado de un recurso, el usuario puede modificar las etiquetas y eludir la intención de las políticas basadas en etiquetas.

Ejemplo de instrucciones de políticas basadas en identidades para clústeres

Los ejemplos que se muestran a continuación muestran las políticas de permisos basadas en identidades que se utilizan para controlar las acciones que se permiten con clústeres de EMR.

> **Important**
>
> La acción `ModifyInstanceGroup` en Amazon EMR no requiere que especifique un ID de clúster. Por ese motivo, denegar esta acción en función de las etiquetas de clúster requiere una consideración adicional. Para obtener más información, consulte Denegar la ModifyInstanceGroup acción en Amazon EMR.

Temas

- Permitir acciones solo en clústeres con valores de etiqueta específicos
- Etiquetado obligatorio de un clúster al crear un clúster
- Permitir acciones en clústeres con una etiqueta específica, con independencia del valor de su etiqueta

Permitir acciones solo en clústeres con valores de etiqueta específicos

Los ejemplos que aparecen a continuación muestran una política que permite a un usuario realizar acciones en función de la etiqueta de clúster *`department`* con el valor *`dev`* y también permite a un usuario etiquetar clústeres con la misma etiqueta. El ejemplo de política muestra cómo denegar los privilegios para etiquetar clústeres de EMR con cualquier otra cosa excepto la misma etiqueta.

En el siguiente ejemplo de política, la condición `StringEquals` intenta hacer coincidir *`dev`* con el valor de la etiqueta *`department`*. Si la etiqueta *`department`* no se ha añadido al clúster o no contiene el valor *`dev`*, la política no se aplica y esta política no permite las acciones. Si no hay otras instrucciones de política que permitan las acciones, el usuario solo puede trabajar con clústeres que tenga esta etiqueta con este valor.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "Stmt12345678901234",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:DescribeCluster",
        "elasticmapreduce:ListSteps",
```

```
        "elasticmapreduce:TerminateJobFlows",
        "elasticmapreduce:SetTerminationProtection",
        "elasticmapreduce:ListInstances",
        "elasticmapreduce:ListInstanceGroups",
        "elasticmapreduce:ListBootstrapActions",
        "elasticmapreduce:DescribeStep"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/department": "dev"
        }
      }
    }
  ]
}
```

También puede especificar varios valores de etiqueta utilizando un operador de condición. Por ejemplo, a fin de permitir que todas las acciones en clústeres donde la etiqueta *department* contiene el valor *dev* o *test*, podría sustituir el bloque de condición en el ejemplo anterior por los siguientes.

```
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/department":["dev", "test"]
        }
      }
```

Etiquetado obligatorio de un clúster al crear un clúster

Como en el ejemplo anterior, la siguiente política de ejemplo busca la misma etiqueta coincidente: el valor *dev* para la etiqueta *department*. Sin embargo, en este ejemplo, la clave de condición `RequestTag` especifica que la política se aplica durante la creación de la etiqueta. Por lo tanto, debe crear un clúster con una etiqueta que coincida con el valor especificado.

Para crear un clúster con una etiqueta, también debe tener permiso para realizar la acción `elasticmapredue:AddTags`. En esta instrucción, la clave de condición

`elasticmapreduce:ResourceTag` garantiza que IAM solo conceda acceso a los recursos de etiquetas con el valor *`dev`* en la etiqueta *`department`*. El elemento `Resource` se utiliza para limitar este permiso a los recursos del clúster.

En el caso de `PassRole` los recursos, debe proporcionar el identificador o alias de la AWS cuenta, el nombre de la función de servicio en la `PassRoleForEMR` declaración y el nombre del perfil de la instancia en la `PassRoleForEC2` declaración. Para obtener más información sobre el formato ARN de IAM, consulte IAM en la Guía del usuario de ARNs IAM.

Para obtener más información sobre cómo hacer coincidir los valores de las claves de etiqueta, consulte `aws:RequestTag/tag-key` en la Guía del usuario de IAM.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "RunJobFlowExplicitlyWithTag",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:RunJobFlow"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "aws:RequestTag/department": "dev"
        }
      }
    },
    {
      "Sid": "AddTagsForDevClusters",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:AddTags"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:*:*:cluster/*"
      ],
      "Condition": {
```

```
        "StringEquals": {
          "elasticmapreduce:ResourceTag/department": "dev"
        }
      }
    },
    {
      "Sid": "PassRoleForEMR",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/Role-Name-With-Path"
      ],
      "Condition": {
        "StringLike": {
          "iam:PassedToService": "elasticmapreduce.amazonaws.com*"
        }
      }
    },
    {
      "Sid": "PassRoleForEC2",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/Role-Name-With-Path"
      ],
      "Condition": {
        "StringLike": {
          "iam:PassedToService": "ec2.amazonaws.com*"
        }
      }
    }
  ]
}
```

Permitir acciones en clústeres con una etiqueta específica, con independencia del valor de su etiqueta

También puede permitir acciones solo en clústeres que tengan una etiqueta particular, con independencia del valor de la etiqueta. Para ello, puede utilizar el operador `Null`. Para obtener más información, consulte Operador de condición para comprobar la existencia de claves de condición en la Guía del usuario de IAM. Por ejemplo, para permitir acciones solo en clústeres de EMR que tengan la etiqueta *`department`*, con independencia del valor que contenga, podría sustituir los bloques Condition del ejemplo anterior por el siguiente. El operador `Null` buscará la presencia de la etiqueta *`department`* en un clúster de EMR. Si la etiqueta existe, la instrucción `Null` se evalúa como falsa, ajustándose a la condición especificada en esta instrucción de política y se permiten las acciones adecuadas.

```
"Condition": {
  "Null": {
    "elasticmapreduce:ResourceTag/department":"false"
  }
}
```

La siguiente instrucción de política permite a un usuario crear un clúster de EMR solo si el clúster tendrá una etiqueta *`department`*, que puede contener cualquier valor. Para el `PassRole` recurso, debe proporcionar el identificador o alias de la AWS cuenta y el nombre de la función de servicio. Para obtener más información sobre el formato ARN de IAM, consulte IAM en la Guía del usuario de ARNs IAM.

Para obtener más información sobre cómo especificar el operador de condición null (“false”), consulte Operador de condición para comprobar la existencia de claves de condición en la Guía del usuario de IAM.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "CreateClusterTagNullCondition",
      "Effect": "Allow",
      "Action": [
```

```
        "elasticmapreduce:RunJobFlow"
      ],
      "Resource": [
        "*"
      ],
      "Condition": {
        "Null": {
          "aws:RequestTag/department": "false"
        }
      }
    },
    {
      "Sid": "AddTagsNullCondition",
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:AddTags"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:*:*:cluster/*"
      ],
      "Condition": {
        "Null": {
          "elasticmapreduce:ResourceTag/department": "false"
        }
      }
    },
    {
      "Sid": "PassRoleForElasticMapReduce",
      "Effect": "Allow",
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/Role-Name-With-Path"
      ],
      "Condition": {
        "StringLike": {
          "iam:PassedToService": "elasticmapreduce.amazonaws.com*"
        }
      }
    },
    {
      "Sid": "PassRoleForEC2",
      "Effect": "Allow",
```

```
      "Action": [
        "iam:PassRole"
      ],
      "Resource": [
        "arn:aws:iam::123456789012:role/Role-Name-With-Path"
      ],
      "Condition": {
        "StringLike": {
          "iam:PassedToService": "ec2.amazonaws.com*"
        }
      }
    }
  ]
}
```

Ejemplo de instrucciones de políticas basadas en identidades para Cuadernos de Amazon EMR

Las instrucciones de políticas de IAM de ejemplo de esta sección muestran escenarios comunes del uso de claves para limitar las acciones permitidas mediante Cuadernos de Amazon EMR. Siempre que no haya ninguna otra política asociada a la entidad principal (usuario) que permita las acciones, las claves de contexto de condición limitan las acciones permitidas tal como se indica.

Example : permitir el acceso únicamente a Cuadernos de Amazon EMR que un usuario cree en función del etiquetado

La instrucción de política de ejemplo mostrada a continuación, cuando se asocia a un rol o a un usuario, permite al usuario trabajar solamente con los cuadernos que haya creado. Esta instrucción de política utiliza la etiqueta predeterminada que se aplica al crear un bloc de notas.

En el ejemplo, el operador de condición `StringEquals` intenta emparejar una variable que representa el ID de usuario del usuario actual (`{aws:userId}`) con el valor de la etiqueta `creatorUserID`. Si la etiqueta `creatorUserID` no se ha añadido al bloc de notas o no contiene el valor del ID del usuario actual, la política no se aplica y esta política no permite las acciones. Si no hay otras instrucciones de política que permitan las acciones, el usuario solo puede trabajar con blocs de notas que tengan este valor para esta etiqueta.

JSON

```
{
  "Version": "2012-10-17",
```

```
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:DescribeEditor",
        "elasticmapreduce:StartEditor",
        "elasticmapreduce:StopEditor",
        "elasticmapreduce:DeleteEditor",
        "elasticmapreduce:OpenEditorInConsole"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/creatorUserId": "${aws:userId}"
        }
      },
      "Sid": "AllowELASTICMAPREDUCEDescribeeditor"
    }
  ]
}
```

Example — Etiquetado de cuadernos obligatorio al crearlos

En este ejemplo, se utiliza la clave de contexto `RequestTag`. La acción `CreateEditor` solo se permite si el usuario no cambia ni elimina la etiqueta `creatorUserID` que se añade de forma predeterminada. La variable ${aws:userId}, especifica el ID de usuario del usuario activo actualmente, que es el valor predeterminado de la etiqueta.

La instrucción de política se puede utilizar para ayudar a garantizar que los usuarios no eliminan la etiqueta `createUserId` ni cambian su valor.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:CreateEditor"
```

```
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:RequestTag/creatorUserId": "${aws:userid}"
        }
      },
      "Sid": "AllowELASTICMAPREDUCECreateeditor"
    }
  ]
}
```

Este ejemplo requiere que el usuario cree el clúster con una etiqueta que tenga la cadena de clave `dept` establecida en uno de los siguientes valores: `datascience`, `analytics`, `operations`.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:CreateEditor"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:RequestTag/dept": [
            "datascience",
            "analytics",
            "operations"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCECreateeditor"
```

```
    }
  ]
}
```

Example — Limitación de la creación de cuadernos a clústeres etiquetados y etiquetas de cuadernos obligatorias

Este ejemplo permite la creación de blocs de notas solo si el bloc de notas se crea con una etiqueta que tiene la cadena de clave `owner` establecida en uno de los valores especificados. Además, el bloc de notas solo se puede crear si el clúster tiene una etiqueta con la cadena de clave `department` establecida en uno de los valores especificados.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:CreateEditor"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:RequestTag/owner": [
            "owner1",
            "owner2",
            "owner3"
          ],
          "elasticmapreduce:ResourceTag/department": [
            "dep1",
            "dep3"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCECreateeditor"
    }
  ]
```

```
}
```

Example — Limitación de la capacidad de iniciar un cuaderno en función de las etiquetas

Este ejemplo limita la capacidad de iniciar blocs de notas únicamente a aquellos blocs de notas que tienen una etiqueta con la cadena de clave `owner` establecida en uno de los valores especificados. Debido a que el elemento `Resource` se utiliza solo para especificar el valor `editor`, la condición no se aplica al clúster y no es necesario etiquetarlo.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:StartEditor"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:editor/*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/owner": [
            "owner1",
            "owner2"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEStarteditor"
    }
  ]
}
```

Este ejemplo es similar a uno anterior. Sin embargo, el límite solo se aplica a los clústeres etiquetados, no a los blocs de notas.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:StartEditor"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:cluster/*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/department": [
            "dep1",
            "dep3"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEStarteditor"
    }
  ]
}
```

En este ejemplo, se utiliza un conjunto diferente de etiquetas de bloc de notas y de clúster. Permite el inicio de un bloc de notas solo si:

- El bloc de notas tiene una etiqueta con la cadena de clave `owner` establecida en cualquiera de los valores especificados

  —y—
- El clúster tiene una etiqueta con la cadena de clave `department` establecida en cualquiera de los valores especificados

JSON

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:StartEditor"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:editor/*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/owner": [
            "user1",
            "user2"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEStarteditorByOwner"
    },
    {
      "Action": [
        "elasticmapreduce:StartEditor"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:cluster/*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/department": [
            "datascience",
            "analytics"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEStarteditorByDepartment"
    }
  ]
}
```

Example — Limitación de la capacidad de abrir el editor de cuadernos en función de las etiquetas

En este ejemplo, se permite la apertura del editor de blocs de notas solo si:

- El bloc de notas tiene una etiqueta con la cadena de clave `owner` establecida en cualquiera de los valores especificados.

  —y—

- El clúster tiene una etiqueta con la cadena de clave `department` establecida en cualquiera de los valores especificados.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:OpenEditorInConsole"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:editor/*"
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/owner": [
            "user1",
            "user2"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEOpeneditorconsoleByOwner"
    },
    {
      "Action": [
        "elasticmapreduce:OpenEditorInConsole"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:cluster/*"
```

```
      ],
      "Condition": {
        "StringEquals": {
          "elasticmapreduce:ResourceTag/department": [
            "datascience",
            "analytics"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEOpeneditorconsoleByDepartment"
    }
  ]
}
```

Denegar la ModifyInstanceGroup acción en Amazon EMR

La ModifyInstanceGroupsacción en Amazon EMR no requiere que proporcione un ID de clúster junto con la acción. En su lugar, puede especificar solo un ID de grupo de instancias. Por este motivo, es posible que una política de denegación aparentemente simple para esta acción basada en el ID de clúster o en una etiqueta de clúster no tenga el efecto deseado. Considere la política de ejemplo siguiente.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Sid": "AllowELASTICMAPREDUCEModifyinstancegroups"
    },
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
```

```
      "Effect": "Deny",
      "Resource": [
        "arn:aws:elasticmapreduce:us-east-1:123456789012:cluster/
j-12345ABCDEFG67"
      ],
      "Sid": "DenyELASTICMAPREDUCEModifyinstancegroups"
    }
  ]
}
```

Si un usuario con esta política adjunta realiza una acción `ModifyInstanceGroup` y especifica solo el ID del grupo de instancias, la política no se aplica. Como la acción está permitida en todos los demás recursos, se realiza correctamente.

Una solución a este problema consiste en adjuntar a la identidad una declaración de política que utilice un NotResourceelemento para denegar cualquier `ModifyInstanceGroup` acción realizada sin un ID de clúster. El siguiente ejemplo de política agrega una declaración de denegación de este tipo para que cualquier solicitud `ModifyInstanceGroups` falle, a menos que se especifique un ID de clúster. Como una identidad debe especificar un ID de clúster con la acción, las instrucciones de denegación basadas en el ID de clúster son, por lo tanto, eficaces.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Sid": "AllowELASTICMAPREDUCEModifyinstancegroups"
    },
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
```

```
      "Effect": "Deny",
      "Resource": [
        "arn:aws:elasticmapreduce:us-east-1:123456789012:cluster/
j-12345ABCDEFG67"
      ],
      "Sid": "DenyELASTICMAPREDUCEModifyinstancegroupsSpecificCluster"
    },
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Effect": "Deny",
      "NotResource": "arn:*:elasticmapreduce:*:*:cluster/*",
      "Sid": "DenyELASTICMAPREDUCEModifyinstancegroupsNonCluster"
    }
  ]
}
```

Se produce un problema similar cuando se quiere denegar la acción `ModifyInstanceGroups` en función del valor asociado a una etiqueta de clúster. La solución es similar. Además de una sentencia de denegación que especifique el valor de la etiqueta, puede agregar una declaración de política que deniegue la acción `ModifyInstanceGroup` si la etiqueta que ha especificado no está presente, independientemente del valor.

El siguiente ejemplo muestra una política que, cuando se adjunta a una identidad, niega la identidad de la acción `ModifyInstanceGroups` a cualquier clúster con la etiqueta `department` establecida en `dev`. Esta declaración solo es efectiva debido a la declaración de denegación, que utiliza la condición `StringNotLike` para denegar la acción, a menos que la etiqueta `department` esté presente.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Effect": "Allow",
```

```
      "Resource": [
        "*"
      ],
      "Sid": "AllowELASTICMAPREDUCEModifyinstancegroups"
    },
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Condition": {
        "StringEquals": {
          "aws:ResourceTag/department": "dev"
        }
      },
      "Effect": "Deny",
      "Resource": [
        "*"
      ],
      "Sid": "DenyELASTICMAPREDUCEModifyinstancegroupsDevDepartment"
    },
    {
      "Action": [
        "elasticmapreduce:ModifyInstanceGroups"
      ],
      "Condition": {
        "StringNotLike": {
          "aws:ResourceTag/department": "?*"
        }
      },
      "Effect": "Deny",
      "Resource": [
        "*"
      ],
      "Sid": "DenyELASTICMAPREDUCEModifyinstancegroupsNoDepartmentTag"
    }
  ]
}
```

Solución de problemas de identidad y acceso de Amazon EMR

Utilice la siguiente información para diagnosticar y solucionar los problemas habituales que pueden surgir cuando se trabaja con Amazon EMR e IAM.

Temas

- No tengo autorización para realizar una acción en Amazon EMR
- No estoy autorizado a realizar lo siguiente: PassRole
- Quiero permitir que personas ajenas a mi AWS cuenta accedan a mis recursos de Amazon EMR

No tengo autorización para realizar una acción en Amazon EMR

Si AWS Management Console le indica que no está autorizado a realizar una acción, debe ponerse en contacto con su administrador para obtener ayuda. Su administrador es la persona que le facilitó su nombre de usuario y contraseña.

En el siguiente ejemplo, el error se produce cuando el usuario `mateojackson` intenta utilizar la consola para consultar los detalles acerca de un recurso ficticio *`my-example-widget`*, pero no tiene los permisos ficticios `EMR:`*`GetWidget`*.

```
User: arn:aws:iam::123456789012:user/mateojackson is not authorized to perform:
 EMR:GetWidget on resource: my-example-widget
```

En este caso, Mateo pide a su administrador que actualice sus políticas de forma que pueda obtener acceso al recurso *`my-example-widget`* mediante la acción `EMR:`*`GetWidget`*.

No estoy autorizado a realizar lo siguiente: PassRole

Si recibe un error que indica que no tiene autorización para llevar a cabo la acción `iam:PassRole`, las políticas se deben actualizar para permitirle pasar un rol a Amazon EMR.

Algunas Servicios de AWS permiten transferir una función existente a ese servicio en lugar de crear una nueva función de servicio o una función vinculada a un servicio. Para ello, debe tener permisos para transferir el rol al servicio.

En el siguiente ejemplo, el error se produce cuando un usuario de IAM denominado `marymajor` intenta utilizar la consola para realizar una acción en Amazon EMR. Sin embargo, la acción requiere que el servicio cuente con permisos que otorguen un rol de servicio. Mary no tiene permisos para transferir el rol al servicio.

```
User: arn:aws:iam::123456789012:user/marymajor is not authorized to perform:
 iam:PassRole
```

En este caso, las políticas de Mary se deben actualizar para permitirle realizar la acción `iam:PassRole`.

Si necesita ayuda, póngase en contacto con su administrador. AWS El gestionador es la persona que le proporcionó las credenciales de inicio de sesión.

Quiero permitir que personas ajenas a mi AWS cuenta accedan a mis recursos de Amazon EMR

Puede crear un rol que los usuarios de otras cuentas o las personas externas a la organización puedan utilizar para acceder a sus recursos. Puede especificar una persona de confianza para que asuma el rol. En el caso de los servicios que admiten políticas basadas en recursos o listas de control de acceso (ACLs), puede utilizar esas políticas para permitir que las personas accedan a sus recursos.

Para obtener más información, consulte lo siguiente:

- Para saber si Amazon EMR admite estas características, consulte Cómo funciona Amazon EMR con IAM.
- Para obtener información sobre cómo proporcionar acceso a los recursos de su Cuentas de AWS propiedad, consulte Proporcionar acceso a un usuario de IAM en otro de su propiedad en la Cuenta de AWS Guía del usuario de IAM.
- Para obtener información sobre cómo proporcionar acceso a tus recursos a terceros Cuentas de AWS, consulta Cómo proporcionar acceso a recursos que Cuentas de AWS son propiedad de terceros en la Guía del usuario de IAM.
- Para obtener información sobre cómo proporcionar acceso mediante una federación de identidades, consulta Proporcionar acceso a usuarios autenticados externamente (identidad federada) en la Guía del usuario de IAM.
- Para conocer sobre la diferencia entre las políticas basadas en roles y en recursos para el acceso entre cuentas, consulte Acceso a recursos entre cuentas en IAM en la Guía del usuario de IAM.

Uso de Amazon S3 Access Grants con Amazon EMR

Información general de S3 Access Grants para Amazon EMR

Con las versiones 6.15.0 y posteriores de Amazon EMR, Amazon S3 Access Grants proporciona una solución de control de acceso escalable que puede utilizar para aumentar el acceso a los datos de Amazon S3 desde Amazon EMR. Si cuenta con una configuración de permisos compleja o amplia de

datos de S3, puede utilizar las concesiones de acceso a fin de escalar los permisos de datos de S3 para usuarios, grupos, roles y aplicaciones de un clúster.

Utilice S3 Access Grants para incrementar el acceso a los datos de Amazon S3, más allá de los permisos que conceden el rol de tiempo de ejecución o los roles de IAM asociados a las identidades con acceso al clúster de EMR. Para obtener más información, consulte Administración del acceso con S3 Access Grants en la Guía del usuario de Amazon S3.

Para conocer los pasos para utilizar S3 Access Grants con otras implementaciones de Amazon EMR, consulte la siguiente documentación:

- Uso de S3 Access Grants con Amazon EMR en EKS
- Uso de S3 Access Grants con Amazon EMR sin servidor

Cómo funciona Amazon EMR con S3 Access Grants

Las versiones 6.15.0 y posteriores de Amazon EMR proporcionan una integración nativa con S3 Access Grants. Puede habilitar S3 Access Grants en Amazon EMR y ejecutar trabajos de Spark. Cuando un trabajo de Spark solicita datos de S3, Amazon S3 proporciona credenciales temporales que se limitan al bucket, al prefijo o al objeto.

A continuación, se ofrece información general de alto nivel sobre cómo Amazon EMR obtiene acceso a los datos protegidos por S3 Access Grants.

1. Un usuario envía un trabajo de Spark de Amazon EMR que utiliza datos almacenados en Amazon S3.
2. Amazon EMR solicita a S3 Access Grants que permita el acceso al bucket, al prefijo o al objeto en nombre de ese usuario.
3. Amazon S3 devuelve credenciales temporales en forma de token AWS Security Token Service (STS) para el usuario. El token tiene el alcance para acceder al bucket, al prefijo o al objeto de S3.
4. Amazon EMR utiliza el token de STS para recuperar datos de S3.
5. Amazon EMR recibe los datos de S3 y devuelve los resultados al usuario.

Consideraciones sobre el uso de S3 Access Grants con Amazon EMR

Tenga en cuenta los siguientes comportamientos y limitaciones cuando utilice S3 Access Grants con Amazon EMR.

Compatibilidad de características

- S3 Access Grants es compatible con las versiones 6.15.0 y posteriores de Amazon EMR.
- Spark es el único motor de consultas compatible cuando se utiliza S3 Access Grants con Amazon EMR.
- Delta Lake y Hudi son los únicos formatos de tabla abierta compatibles cuando se utiliza S3 Access Grants con Amazon EMR.
- Las siguientes capacidades de Amazon EMR no son compatibles para su uso con S3 Access Grants:
  - Tablas de Apache Iceberg
  - Autenticación nativa de LDAP
  - Autenticación nativa de Apache Ranger
  - AWS CLI solicitudes a Amazon S3 que utilizan funciones de IAM
  - Acceso a S3 mediante el protocolo de código abierto de S3A
- La opción `fallbackToIAM` no es compatible con los clústeres de EMR que utilizan la propagación de identidades de confianza con IAM Identity Center.
- S3 Access Grants with solo AWS Lake Formation es compatible con los clústeres de Amazon EMR que se ejecutan en Amazon. EC2

Consideraciones sobre el comportamiento

- La integración nativa de Apache Ranger con Amazon EMR incluye una funcionalidad congruente con S3 Access Grants como parte del complemento Apache Ranger S3 de EMRFS. Si utiliza Apache Ranger para un control de acceso detallado (FGAC), recomendamos utilizar ese complemento en lugar de S3 Access Grants.
- Amazon EMR proporciona una caché de credenciales en EMRFS para garantizar que un usuario no necesite realizar solicitudes repetidas de las mismas credenciales en un trabajo de Spark. Por lo tanto, Amazon EMR siempre solicita el privilegio de nivel predeterminado cuando solicita credenciales. Para obtener más información, consulte Solicitud de acceso a datos de S3 en la Guía del usuario de Amazon S3.

- En el caso de que un usuario realice una acción que no sea compatible con S3 Access Grants, Amazon EMR está configurado para utilizar el rol de IAM que se especificó para la ejecución del trabajo. Para obtener más información, consulte Alternativa de roles de IAM.

Lanzamiento de un clúster de Amazon EMR con S3 Access Grants

En esta sección se describe cómo lanzar un clúster de EMR que se ejecute en Amazon EC2 y utilice S3 Access Grants para administrar el acceso a los datos en Amazon S3. Para conocer los pasos para utilizar S3 Access Grants con otras implementaciones de Amazon EMR, consulte la siguiente documentación:

- Uso de S3 Access Grants con Amazon EMR en EKS
- Uso de S3 Access Grants con EMR sin servidor

Siga los siguientes pasos para lanzar un clúster de EMR que se ejecute en Amazon EC2 y utilice S3 Access Grants para administrar el acceso a los datos en Amazon S3.

1. Configure un rol de ejecución de trabajos para el clúster de EMR. Incluya los permisos de IAM necesarios para ejecutar los trabajos de Spark, `s3:GetDataAccess` y `s3:GetAccessGrantsInstanceForPrefix`:

```
{
    "Effect": "Allow",
    "Action": [
    "s3:GetDataAccess",
    "s3:GetAccessGrantsInstanceForPrefix"
    ],
    "Resource": [     //LIST ALL INSTANCE ARNS THAT THE ROLE IS ALLOWED TO QUERY
         "arn:aws_partition:s3:Region:account-id1:access-grants/default",
         "arn:aws_partition:s3:Region:account-id2:access-grants/default"
    ]
}
```

> **Note**
>
> Con Amazon EMR, S3 Access Grants aumenta los permisos que se configuran en los roles de IAM. Si los roles de IAM que especifica para la ejecución de trabajos contienen

permisos para acceder directamente a S3, es posible que los usuarios puedan acceder a más datos además de los que usted define en S3 Access Grants.

2. A continuación, utilice AWS CLI para crear un clúster con Amazon EMR 6.15 o superior y la `emrfs-site` clasificación para habilitar S3 Access Grants, similar al ejemplo siguiente:

```
aws emr create-cluster
  --release-label emr-6.15.0 \
  --instance-count 3 \
  --instance-type m5.xlarge \
  --configurations '[{"Classification":"emrfs-site",
 "Properties":{"fs.s3.s3AccessGrants.enabled":"true",
 "fs.s3.s3AccessGrants.fallbackToIAM":"false"}}]'
```

S3 Access Grants con AWS Lake Formation

Si utiliza Amazon EMR con la integración de AWS Lake Formation, puede utilizar Amazon S3 Access Grants para el acceso directo o tabular a los datos de Amazon S3.

> **Note**
>
> S3 Access Grants with solo AWS Lake Formation es compatible con los clústeres de Amazon EMR que se ejecutan en Amazon. EC2

Acceso directo

El acceso directo implica todas las llamadas para acceder a los datos de S3 que no invoquen la API del servicio AWS Glue que Lake Formation utiliza como metaalmacén con Amazon EMR, por ejemplo, para llamar a: `spark.read`

```
spark.read.csv("s3://...")
```

Cuando utiliza S3 Access Grants con AWS Lake Formation Amazon EMR, todos los patrones de acceso directo pasan por S3 Access Grants para obtener credenciales de S3 temporales.

Acceso tabular

El acceso tabular se produce cuando Lake Formation invoca la API del almacén de metadatos para acceder a su ubicación de S3, por ejemplo, para consultar datos de tablas:

```
spark.sql("select * from test_tbl")
```

Cuando utiliza S3 Access Grants con AWS Lake Formation Amazon EMR, todos los patrones de acceso tabulares pasan por Lake Formation.

Alternativa de roles de IAM

Si un usuario intenta realizar una acción que no es compatible con S3 Access Grants, Amazon EMR está configurado para utilizar de forma predeterminada el rol de IAM que se especificó para la ejecución de trabajos cuando la configuración `fallbackToIAM` es `true`. Esto permite a los usuarios recurrir al rol de ejecución de trabajos para proporcionar credenciales de acceso a S3 en situaciones que S3 Access Grants no cubre.

Si `fallbackToIAM` está habilitado, los usuarios pueden acceder a los datos que permite Access Grant. Si no hay un token de S3 Access Grants para los datos de destino, Amazon EMR comprueba el permiso en su rol de ejecución de trabajos.

> **Note**
>
> Recomendamos probar los permisos de acceso con la configuración `fallbackToIAM` habilitada, incluso si planea deshabilitar la opción para las cargas de trabajo de producción. Con los trabajos de Spark, hay otras formas en las que los usuarios pueden acceder a todos los conjuntos de permisos con sus credenciales de IAM. Cuando se habilitan en los clústeres de EMR, las concesiones de S3 permiten que los trabajos de Spark accedan a las ubicaciones de S3. Debe asegurarse de proteger estas ubicaciones de S3 del acceso externo a EMRFS. Por ejemplo, debe proteger las ubicaciones de S3 del acceso de los clientes de S3 que se utilizan en las computadoras portátiles o de las aplicaciones que no son compatibles con S3 Access Grants, como Hive o Presto.

Autenticación en nodos de clúster de Amazon EMR

Los clientes SSH pueden usar un par de EC2 claves de Amazon para autenticarse en las instancias del clúster. De forma alternativa, con las versiones 5.10.0 y posteriores de Amazon EMR, puede configurar Kerberos para autenticar a los usuarios y las conexiones SSH con el nodo principal. Además, con las versiones 5.12.0 y posteriores de Amazon EMR, puede autenticarse con LDAP.

Temas

- Utilice un par de EC2 claves para las credenciales SSH de Amazon EMR
- Uso de Kerberos para la autenticación con Amazon EMR
- Uso de servidores de Active Directory o LDAP para la autenticación con Amazon EMR

Utilice un par de EC2 claves para las credenciales SSH de Amazon EMR

Los nodos del clúster de Amazon EMR se ejecutan en instancias de Amazon EC2 . Puede conectarse a los nodos del clúster de la misma manera que a las EC2 instancias de Amazon. Puedes usar Amazon EC2 para crear un par de claves o puedes importar un par de claves. Al crear un clúster, puedes especificar el par de EC2 claves de Amazon que se utilizará para las conexiones SSH a todas las instancias del clúster. También puede crear un clúster sin un par de claves. Esto se hace con clústeres transitorios que se inician, ejecutan pasos, y luego se terminan de forma automática.

El cliente SSH que se utiliza para conectar al clúster necesita utilizar el archivo de clave privada asociado a este par de claves. Se trata de un archivo.pem para clientes SSH que utilizan Linux, Unix y MacOS. Debe establecer permisos para que solo el propietario de la clave tenga permiso para acceder al archivo. Para los clientes SSH que utilizan Windows, es un archivo .ppk, que normalmente se crea a partir del archivo .pem.

- Para obtener más información sobre la creación de un par de EC2 claves de Amazon, consulta los pares de EC2 claves de Amazon en la Guía del EC2 usuario de Amazon.
- Para obtener instrucciones sobre cómo usar Pu TTYgen para crear un archivo.ppk a partir de un archivo.pem, consulta Cómo convertir tu clave privada con Pu en TTYgen la Guía del usuario de Amazon. EC2
- Para obtener más información sobre cómo configurar los permisos del archivo.pem y cómo conectarse al nodo principal de un clúster de EMR mediante diferentes métodos, incluidos `ssh` Linux o macOS, PuTTY desde Windows o AWS CLI desde cualquier sistema operativo compatible, consulte. Conexión al nodo principal del clúster de Amazon EMR mediante SSH

Uso de Kerberos para la autenticación con Amazon EMR

Las versiones 5.10.0 y posteriores de Amazon EMR son compatibles con Kerberos. Kerberos es un protocolo de autenticación de redes que utiliza la criptografía de clave secreta para proporcionar una

autenticación sólida, de forma que las contraseñas u otras credenciales no se envíen a través de la red en un formato no cifrado.

En Kerberos, los servicios y los usuarios que necesitan autenticarse se conocen como entidades principales. Las entidades principales existen dentro de un ámbito de Kerberos. En el ámbito, un servidor de Kerberos conocido como centro de distribución de claves (KDC) proporciona los medios para que se autentiquen las entidades principales. El KDC lo hace mediante la emisión de tickets para la autenticación. El KDC mantiene una base de datos de las entidades principales dentro de su ámbito, sus contraseñas y otros datos administrativos sobre cada una de las entidades principales. Un KDC también puede aceptar credenciales de autenticación de entidades principales de otros ámbitos, lo que se conoce como una confianza entre ámbitos. Además, un clúster de EMR puede utilizar un KDC externo para autenticar principales.

Una forma común de establecer una relación de confianza entre ámbitos o de utilizar un KDC externo es autenticar a los usuarios desde un dominio de Active Directory. Esto permite a los usuarios obtener acceso a un clúster de EMR con su cuenta de dominio cuando utilizan SSH para conectarse a un clúster o trabajan con aplicaciones de macrodatos.

Si utiliza autenticación de Kerberos, Amazon EMR configura Kerberos para las aplicaciones, componentes y subsistemas que instala en el clúster, de forma que se autentiquen entre sí.

> **Important**
>
> Amazon EMR no se admite AWS Directory Service for Microsoft Active Directory en un fideicomiso entre dominios ni como un KDC externo.

Antes de configurar Kerberos con Amazon EMR, le recomendamos que se familiarice con los conceptos de Kerberos, los servicios que se ejecutan en un KDC y las herramientas de administración de servicios de Kerberos. Para obtener más información, consulte la documentación de Kerberos del MIT, que publica el consorcio Kerberos.

Temas

- Aplicaciones compatibles con Amazon EMR
- Opciones de la arquitectura de Kerberos con Amazon EMR
- Configuración de Kerberos en Amazon EMR
- Uso de SSH para conectarse a clústeres que utilizan Kerberos con Amazon EMR
- Tutorial: Configuración de un KDC dedicado del clúster con Amazon EMR

- Tutorial: Configuración de una relación de confianza entre ámbitos con un dominio de Active Directory

Aplicaciones compatibles con Amazon EMR

Dentro de un clúster de EMR, las entidades principales de Kerberos son los servicios de aplicaciones de macrodatos y los subsistemas que se ejecutan en todos los nodos del clúster. Amazon EMR puede configurar las aplicaciones y componentes que se indican a continuación para utilizar Kerberos. Cada aplicación tiene una entidad principal de Kerberos asociada.

Amazon EMR no admite las relaciones de confianza entre ámbitos con AWS Directory Service for Microsoft Active Directory.

Amazon EMR solo configura las características de autenticación de Kerberos de código abierto para las aplicaciones y los componentes indicados a continuación. Cualquier otra aplicación instalada no utiliza Kerberos, lo que puede provocar la incapacidad para comunicarse con los componentes que utilizan Kerberos y generar errores en la aplicación. Las aplicaciones y componentes que no utilizan Kerberos no tienen habilitada la autenticación. Las aplicaciones y los componentes compatibles pueden variar según las distintas versiones de Amazon EMR.

La interfaz de usuario de Livy es la única interfaz de usuario web alojada en el clúster que está kerberizado.

- Hadoop MapReduce
- HBase
- HCatalog
- HDFS
- Hive
  - No active Hive con la autenticación LDAP. Esto puede provocar problemas al comunicarse con el YARN que utiliza Kerberos.
- Hue
  - Hue la autenticación de usuarios no se establece automáticamente y se pueden configurar mediante la configuración de la API.
  - El servidor de Hue utiliza Kerberos. El front-end (la UI) de Hue no está configurado para la autenticación. La autenticación LDAP se puede configurar para la UI de Hue.
- Livy

  - La suplantación de Livy con clústeres kerberizados es compatible con las versiones 5.22.0 y posteriores de Amazon EMR.
- Oozie
- Phoenix
- Presto
  - Presto admite la autenticación de Kerberos en las versiones 6.9.0 y posteriores de Amazon EMR.
  - Si quiere usar la autenticación de Kerberos para Presto, debe habilitar el cifrado en tránsito.
- Spark
- Tez
- Trino
  - Trino admite la autenticación de Kerberos en las versiones 6.11.0 y posteriores de Amazon EMR.
  - Si quiere usar la autenticación de Kerberos para Trino, debe habilitar el cifrado en tránsito.
- YARN
- Zeppelin
  - Zeppelin solo está configurado para utilizar Kerberos con el intérprete Spark. No está configurado para otros intérpretes.
  - Los intérpretes de Zeppelin kerberizado que no sean Spark no admiten la suplantación de identidad de usuario.
- ZooKeeper
  - Zookeeper cliente no es compatible.

Opciones de la arquitectura de Kerberos con Amazon EMR

Cuando se utiliza Kerberos con Amazon EMR, puede elegir entre las arquitecturas que se indican en esta sección. Independientemente de la arquitectura que elija, para configurar Kerberos hay que seguir los mismos pasos. Debe crear una configuración de seguridad, especificar la configuración de seguridad y las opciones de Kerberos específicas del clúster compatibles al crear el clúster y crear directorios de HDFS para los usuarios de Linux en el clúster que coincidan con las entidades principales de usuarios en el KDC. Para leer una explicación sobre las opciones de configuración y configuraciones de ejemplo de cada arquitectura, consulte Configuración de Kerberos en Amazon EMR.

KDC dedicado del clúster (KDC en nodo principal)

Esta configuración está disponible con las versiones 5.10.0 y posteriores de Amazon EMR.

Ventajas

- Amazon EMR tiene la plena propiedad del KDC.
- El KDC en el clúster de EMR es independiente de las implementaciones de KDC centralizadas, como Microsoft Active Directory o AWS Managed Microsoft AD.
- El impacto en el rendimiento es mínimo, ya que el KDC administra la autenticación solo para los nodos locales del clúster.
- Opcionalmente, otros clústeres que utilizan Kerberos pueden hacer referencia al KDC como un KDC externo. Para obtener más información, consulte KDC externo: nodo principal en un clúster diferente.

Consideraciones y limitaciones

- Los clústeres que utilizan Kerberos no se pueden autenticarse entre sí, por lo que las aplicaciones no pueden interactuar. Si las aplicaciones de clústeres tienen que interactuar, debe establecer una relación de confianza entre ámbitos entre los clústeres o configurar un clúster como el KDC externo para otros clústeres. Si se establece una confianza entre dominios, estos KDCs deben tener diferentes dominios de Kerberos.
- Debe crear los usuarios de Linux en la EC2 instancia del nodo principal que correspondan a los usuarios principales del KDC, junto con los directorios HDFS de cada usuario.
- Los usuarios principales deben usar un archivo de clave EC2 privada y `kinit` credenciales para conectarse al clúster mediante SSH.

Relación de confianza entre ámbitos

En esta configuración, las entidades principales (normalmente usuarios) de otro ámbito de Kerberos se autentican para los componentes de la aplicación en un clúster de EMR que utiliza Kerberos y que tiene su propio KDC. El KDC del nodo principal establece una relación de confianza con otro KDC mediante un principal entre dominios que existe en ambos. KDCs El nombre y la contraseña de la entidad principal deben coincidir exactamente en cada KDC. Las relaciones de confianza entre ámbitos son más comunes con las implementaciones de Active Directory, tal y como se muestra en

el siguiente diagrama. También se admiten relaciones de confianza entre ámbitos con un KDC de MIT externo o un KDC en otro clúster de Amazon EMR.

Ventajas

- El clúster de EMR en el que está instalado el KDC mantiene la plena propiedad del KDC.
- Con Active Directory, Amazon EMR crea automáticamente usuarios de Linux que se corresponden con las entidades principales de usuario del KDC. Aun así, debe crear directorios de HDFS para cada usuario. Además, los usuarios principales del dominio de Active Directory pueden acceder a los clústeres kerberizados mediante `kinit` credenciales, sin el archivo de clave privada. EC2 Así se elimina la necesidad de compartir el archivo de clave privada entre usuarios de clústeres.
- Dado que cada KDC del clúster administra la autenticación de los nodos del clúster, se minimizan los efectos de la latencia de la red y la sobrecarga de procesamiento de un gran número de nodos entre clústeres.

Consideraciones y limitaciones

- Si va a establecer una relación de confianza con un ámbito de Active Directory, debe proporcionar un nombre de usuario y contraseña de Active Directory con permisos para unir entidades principales al dominio al crear el clúster.
- Las relaciones de confianza entre ámbitos no se pueden establecer entre ámbitos de Kerberos con el mismo nombre.
- Las relaciones de confianza entre ámbitos deben establecerse de forma explícita. Por ejemplo, si el Clúster A y B establecen una relación de confianza entre ámbitos con un KDC, no confían intrínsecamente uno en el otro y sus aplicaciones no pueden autenticarse entre sí para interaccionar.
- KDCs deben mantenerse de forma independiente y coordinada para que las credenciales de los usuarios principales coincidan con precisión.

KDC externo

Las configuraciones con un KDC externo son compatibles con las versiones 5.20.0 y posteriores de Amazon EMR.

- KDC externo: KDC de MIT
- KDC externo: nodo principal en un clúster diferente

- KDC externo: KDC de clúster en un clúster diferente con una relación de confianza entre ámbitos de Active Directory

KDC externo: KDC de MIT

Esta configuración permite que uno o varios clústeres de EMR utilicen entidades principales definidas y mantenidos en un servidor de KDC de MIT.

Ventajas

- La administración de principales se consolida en un solo KDC.
- Varios clústeres pueden utilizar el mismo KDC en el mismo ámbito de Kerberos. Para obtener más información, consulte Requisitos para utilizar varios clústeres con el mismo KDC.
- El nodo principal en un clúster que utiliza Kerberos no tiene la carga de rendimiento asociada con el mantenimiento del KDC.

Consideraciones y limitaciones

- Debe crear usuarios de Linux en la EC2 instancia del nodo principal de cada clúster kerberizado que correspondan a los usuarios principales de KDC, junto con los directorios HDFS de cada usuario.
- Los usuarios principales deben usar un archivo de claves EC2 privadas y `kinit` credenciales para conectarse a los clústeres Kerberizados mediante SSH.
- Cada nodo de los clústeres de EMR que utilizan Kerberos debe tener una ruta de red al KDC.
- Cada nodo en los clústeres que utilizan Kerberos impone un carga de autenticación en el KDC externo, por lo que la configuración del KDC afecta al rendimiento del clúster. Cuando configure el hardware del servidor de KDC, tenga en cuenta el número máximo de nodos de Amazon EMR que se pueden admitir de forma simultánea.
- El rendimiento del clúster depende de la latencia de red entre los nodos en los clústeres que utilizan Kerberos y el KDC.
- La solución de problemas puede ser más difícil debido a las interdependencias.

KDC externo: nodo principal en un clúster diferente

Esta configuración es casi idéntica a la implementación de KDC de MIT externo anterior, salvo que el KDC está en el nodo principal de un clúster de EMR. Para obtener más información, consulte KDC dedicado del clúster (KDC en nodo principal) y Tutorial: Configuración de una relación de confianza entre ámbitos con un dominio de Active Directory.

Ventajas

- La administración de principales se consolida en un solo KDC.
- Varios clústeres pueden utilizar el mismo KDC en el mismo ámbito de Kerberos. Para obtener más información, consulte Requisitos para utilizar varios clústeres con el mismo KDC.

Consideraciones y limitaciones

- Debe crear usuarios de Linux en la EC2 instancia del nodo principal de cada clúster kerberizado que correspondan a los usuarios principales de KDC, junto con los directorios HDFS de cada usuario.
- Los usuarios principales deben usar un archivo de claves EC2 privadas y `kinit` credenciales para conectarse a los clústeres Kerberizados mediante SSH.
- Cada nodo de cada clúster de EMR debe tener una ruta de red al KDC.
- Cada nodo de Amazon EMR de los clústeres que utilizan Kerberos impone un carga de autenticación en el KDC externo, por lo que la configuración del KDC afecta al rendimiento del clúster. Cuando configure el hardware del servidor de KDC, tenga en cuenta el número máximo de nodos de Amazon EMR que se pueden admitir de forma simultánea.
- El rendimiento del clúster depende de la latencia de red entre los nodos de clústeres y el KDC.
- La solución de problemas puede ser más difícil debido a las interdependencias.

KDC externo: KDC de clúster en un clúster diferente con una relación de confianza entre ámbitos de Active Directory

En esta configuración, primero se crea un clúster con un KDC dedicado del clúster que tiene una relación de confianza entre ámbitos unidireccional con Active Directory. Para ver un tutorial detallado, consulte Tutorial: Configuración de una relación de confianza entre ámbitos con un dominio de Active Directory. A continuación, lance clústeres adicionales, haciendo referencia al KDC del clúster que

tiene la relación de confianza con un KDC externo. Para ver un ejemplo, consulta KDC externo del clúster con una relación de confianza entre ámbitos de Active Directory. De este modo, cada clúster de Amazon EMR puede utilizar el KDC externo para autenticar entidades principales definidas y mantenidas en un dominio de Microsoft Active Directory.

Ventajas

- La administración de entidades principales se consolida en el dominio de Active Directory.
- Amazon EMR se incorpora al ámbito de Active Directory, lo que elimina la necesidad de crear usuarios de Linux que se correspondan con usuarios de Active Directory. Aun así, debe crear directorios de HDFS para cada usuario.
- Varios clústeres pueden utilizar el mismo KDC en el mismo ámbito de Kerberos. Para obtener más información, consulte Requisitos para utilizar varios clústeres con el mismo KDC.
- Los usuarios principales del dominio de Active Directory pueden acceder a los clústeres Kerberizados mediante `kinit` credenciales, sin el archivo de clave privada. EC2 Así se elimina la necesidad de compartir el archivo de clave privada entre usuarios de clústeres.
- Solo un nodo principal de Amazon EMR tiene la carga del mantenimiento del KDC y solo ese clúster debe crearse con credenciales de Active Directory para la relación de confianza entre ámbitos entre el KDC y Active Directory.

Consideraciones y limitaciones

- Cada nodo de cada clúster de EMR debe tener una ruta de red al KDC y el controlador de dominio de Active Directory.
- Cada nodo de Amazon EMR impone un carga de autenticación en el KDC externo, por lo que la configuración del KDC afecta al rendimiento del clúster. Cuando configure el hardware del servidor de KDC, tenga en cuenta el número máximo de nodos de Amazon EMR que se pueden admitir de forma simultánea.
- El rendimiento del clúster depende de la latencia de red entre los nodos de los clústeres y el servidor de KDC.
- La solución de problemas puede ser más difícil debido a las interdependencias.

Requisitos para utilizar varios clústeres con el mismo KDC

Varios clústeres pueden utilizar el mismo KDC en el mismo ámbito de Kerberos. Sin embargo, si los clústeres se ejecutan simultáneamente, es posible que se produzcan errores si utilizan nombres de ServicePrincipal Kerberos que entren en conflicto.

Si tiene varios clústeres simultáneos con el mismo KDC externo, asegúrese de que los clústeres utilicen distintos ámbitos de Kerberos. Si los clústeres deben usar el mismo dominio de Kerberos, asegúrese de que estén en subredes diferentes y de que sus rangos de CIDR no se superpongan.

Configuración de Kerberos en Amazon EMR

En esta sección, se proporcionan detalles y ejemplos de la configuración de Kerberos con arquitecturas comunes. Independientemente de la arquitectura que elija, los aspectos básicos de la configuración son los mismos y se realizan en tres pasos. Si utiliza un KDC externo o configura una relación de confianza entre ámbitos, debe asegurarse de que todos los nodos de un clúster tengan una ruta de red al KDC externo, incluida la configuración de los grupos de seguridad aplicables para permitir el tráfico de entrada y salida de Kerberos.

Paso 1: Cree una configuración de seguridad con propiedades de Kerberos

La configuración de seguridad especifica detalles sobre el KDC de Kerberos y permite reutilizar la configuración de Kerberos cada vez que se crea un clúster. Puede crear una configuración de seguridad mediante la consola de Amazon EMR AWS CLI, la o la API de EMR. La configuración de seguridad también pueden contener otras opciones de seguridad, como, por ejemplo, el cifrado. Para obtener más información acerca de la creación y especificación de las configuraciones de seguridad cuando se crea un clúster, consulte Uso de configuraciones de seguridad para configurar la seguridad del clúster de Amazon EMR. Para obtener más información acerca de las propiedades de Kerberos en una configuración de seguridad, consulte Configuración de Kerberos para las configuraciones de seguridad.

Paso 2: Cree un clúster y especifique los atributos de Kerberos específicos del clúster

Al crear un clúster, especifique la configuración de seguridad de Kerberos junto con las opciones de Kerberos específicas del clúster. Cuando utilice la consola de Amazon EMR, solo están disponibles las opciones de Kerberos compatibles con la configuración de seguridad especificada. Cuando utilice la AWS CLI API de Amazon EMR, asegúrese de especificar las opciones de Kerberos compatibles con la configuración de seguridad especificada. Por ejemplo, si especifica la contraseña de la entidad principal para una relación de confianza entre ámbitos al crear un clúster mediante la CLI y la configuración de seguridad especificada no tiene parámetros de relación de confianza entre

ámbitos, se produce un error. Para obtener más información, consulte Configuración de Kerberos para clústeres.

Paso 3: configure el nodo principal del clúster

En función de los requisitos de su arquitectura e implementación, es posible que sea necesario realizar una configuración adicional en el clúster. Puede hacer esto después de crearla o mediante pasos o acciones de arranque durante el proceso de creación.

Para cada usuario autenticado de Kerberos que se conecte al clúster utilizando SSH, debe asegurarse de que se crean cuentas de Linux que se corresponden con el usuario de Kerberos. Si el controlador de dominio de Active Directory proporciona las entidades principales de usuario, ya sea como el KDC externo o a través de una relación de confianza entre ámbitos, Amazon EMR crea cuentas de usuario de Linux de forma automática. Si no se utiliza Active Directory, debe crear entidades principales para cada usuario que se correspondan con su usuario de Linux. Para obtener más información, consulte Configuración de un clúster de Amazon EMR para conexiones SSH y usuarios de HDFS autenticados en Kerberos.

Cada usuario también debe tener un directorio de usuarios de HDFS propio que debe crear usted. Además, SSH debe estar configurado con GSSAPI habilitado para permitir conexiones de los usuarios autenticados por Kerberos. GSSAPI debe estar habilitado en el nodo principal y se debe configurar la aplicación SSH del cliente para que utilice GSSAPI. Para obtener más información, consulte Configuración de un clúster de Amazon EMR para conexiones SSH y usuarios de HDFS autenticados en Kerberos.

Configuración de seguridad y configuración de clústeres para Kerberos en Amazon EMR

Al crear un clúster que utiliza Kerberos, debe especificar la configuración de seguridad junto con los atributos de Kerberos específicos para el clúster. No puede especificar un conjunto sin el otro, o se producirá un error.

En este tema, se ofrece información general de los parámetros de configuración disponibles para Kerberos al crear una configuración de seguridad y un clúster. Además, se proporcionan ejemplos de CLI para crear configuraciones de seguridad compatibles y clústeres para arquitecturas comunes.

Configuración de Kerberos para las configuraciones de seguridad

Puede crear una configuración de seguridad que especifique los atributos de Kerberos mediante la consola Amazon EMR, AWS CLI la o la API EMR. La configuración de seguridad también pueden contener otras opciones de seguridad, como, por ejemplo, el cifrado. Para obtener más información, consulte Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI.

Utilice las siguientes referencias para conocer las opciones de configuración de seguridad disponibles para la arquitectura Kerberos que elija. Se muestra la configuración de la consola de Amazon EMR. Para ver las opciones de la CLI correspondientes, consulte Especificar la configuración de Kerberos mediante el AWS CLI o Ejemplos de configuraciones.

Parámetro		Descripción
Kerberos		Especifica que Kerberos está habilitado para los clústeres que utilizan esta configuración de seguridad . Si un clúster usa esta configuración de seguridad , también debe tener la configuración de Kerberos especificada o se producirá un error.
Proveedor	KDC dedicado del clúster	Especifica que Amazon EMR crea un KDC en el nodo principal de cualquier clúster que utilice esta configura ción de seguridad. Al crear el clúster, debe especific ar el nombre del dominio y la contraseña de administr ador del KDC. Si es necesario, puede hacer referencia a este KDC desde otros clústeres. Cree esos clústeres con una configuración de seguridad diferente, especifique un KDC externo y utilice el nombre de dominio y la contraseña de administrador del KDC que especifique para el KDC dedicado al clúster.
	KDC externo	Solo está disponible en la versión de Amazon EMR 5.20.0 y posteriores. Especifica que los clústeres que utilizan esta configuración de seguridad autentican las entidades principales de Kerberos mediante un servidor de KDC externo al clúster. No se crea un KDC en el clúster. Al crear el clúster, debe especificar el nombre de dominio y la contraseña de administrador del KDC para el KDC externo.
Ciclo de vida del ticket		Opcional. Especifica el período durante el que un ticket de Kerberos emitido por el KDC es válido en los clústeres que utilizan esta configuración de seguridad.

Parámetro		Descripción
		El ciclo de vida de los tickets es limitado por motivos de seguridad. Las aplicaciones de clúster y los servicios renuevan automáticamente los tickets después de que expiren. Los usuarios que se conecten al clúster mediante SSH utilizando las credenciales de Kerberos tienen que ejecutar el comando `kinit` desde la línea de comandos del nodo principal para la renovación tras la expiración de un ticket.
Relación de confianza entre ámbitos		Especifica una relación de confianza entre ámbitos entre un KDC dedicado a un clúster en los clústeres que utilizan esta configuración de seguridad y un KDC situado en un ámbito de Kerberos diferente. Las entidades principales (normalmente los usuarios) de otro ámbito se autentican en los clústeres que utilizan esta configuración. Se requiere una configura ción adicional en el otro ámbito de Kerberos. Para obtener más información, consulte Tutorial: Configura ción de una relación de confianza entre ámbitos con un dominio de Active Directory.
Propiedad es de la relación de confianza entre ámbitos	Ámbito	Especifica el nombre de ámbito de Kerberos del otro ámbito en la relación de confianza. Por convención, los nombres de ámbito de Kerberos son los mismos que los nombres de dominio, pero en mayúsculas.
	Dominio	Especifica el nombre de dominio del otro ámbito en la relación de confianza.

Parámetro		Descripción
	Servidor de administración	Especifica el nombre de dominio completo (FQDN) o dirección IP del servidor de administración del otro ámbito de la relación de confianza. El servidor de administración y el servidor de KDC suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos para comunicarse. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 749. También se puede especificar el puerto (por ejemplo, `domain.example.com:749`).
	Servidor de KDC	Especifica el nombre de dominio completo (FQDN) o dirección IP del servidor de KDC del otro ámbito de la relación de confianza. El servidor de KDC y el servidor de administración suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 88. También se puede especificar el puerto (por ejemplo, `domain.example.com:88`).
KDC externo		Especifica que el clúster utiliza el KDC externo del clúster.

Parámetro			Descripción
Propiedades de KDC externo	Servidor de administración		Especifica el nombre de dominio completo (FQDN) o la dirección IP del servidor de administración externo. El servidor de administración y el servidor de KDC suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos para comunicarse. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 749. También se puede especificar el puerto (por ejemplo, `domain.example.com:749`).
	Servidor de KDC		Especifica el nombre de dominio completo (FQDN) del servidor de KDC externo. El servidor de KDC y el servidor de administración suelen ejecutarse en el mismo equipo con el mismo FQDN, pero utilizan puertos distintos. Si no se especifica ningún puerto, se usa el puerto predeterminado de Kerberos: el 88. También se puede especificar el puerto (por ejemplo, `domain.example.com:88`).
	Integración de Active Directory		Especifica que la autenticación de la entidad principal de Kerberos está integrada en un dominio de Microsoft Active Directory.
	Propiedades de la integración de Active Directory	Ámbito de Active Directory	Especifica el nombre de ámbito de Kerberos del dominio de Active Directory. Por convención, los nombres de ámbito de Kerberos suelen ser los mismos que los nombres de dominio, pero en mayúsculas.
		Dominio de Active Directory	Especifica el nombre de dominio de Active Directory.

Parámetro			Descripción
		Servidor de Active Directory	Especifica el nombre de dominio completo (FQDN) del controlador de dominio de Microsoft Active Directory.

Configuración de Kerberos para clústeres

Puede especificar la configuración de Kerberos al crear un clúster mediante la consola Amazon EMR, AWS CLI la o la API EMR.

Utilice las siguientes referencias para conocer las opciones de configuración de clústeres disponibles para la arquitectura Kerberos que elija. Se muestra la configuración de la consola de Amazon EMR. Para ver las opciones de la CLI correspondientes, consulte Ejemplos de configuraciones.

Parámetro	Descripción
Ámbito	El nombre del ámbito de Kerberos para el clúster. La convención de Kerberos consiste en establecerlo igual que el nombre del dominio, pero en mayúsculas. Por ejemplo, para el dominio `ec2.internal` , se utiliza `EC2.INTERNAL` como nombre de ámbito.
Contraseña de administración de KDC	La contraseña utilizada en el clúster `kadmin` o `kadmin.local` . Son interfaces de línea de comandos para el sistema de administración de Kerberos V5, que mantiene las entidades principales, las políticas de contraseñas y las tablas de claves de Kerberos para el clúster.
Contraseña de la entidad principal de confianza entre ámbitos (opcional)	Es necesaria para establecer una confianza entre ámbitos. La contraseña de la entidad principal de confianza entre ámbitos, que debe

Parámetro	Descripción
	ser idéntica en los distintos ámbitos. Use una contraseña segura.
Usuario de incorporación al dominio de Active Directory (opcional)	Obligatorio cuando se utiliza Active Directory en una relación de confianza entre ámbitos. Se trata de un nombre de inicio de sesión de usuario de una cuenta de Active Directory con permisos para incorporar equipos al dominio. Amazon EMR utiliza esta identidad para incorporar el clúster al dominio. Para obtener más información, consulte the section called "Paso 3: adición de cuentas al dominio para el clúster de EMR".
Contraseña de incorporación al dominio de Active Directory (opcional)	La contraseña del usuario de incorporación a un dominio de Active Directory. Para obtener más información, consulte the section called "Paso 3: adición de cuentas al dominio para el clúster de EMR".

Ejemplos de configuraciones

Los siguientes ejemplos muestran las configuraciones de seguridad y las configuraciones de clúster para escenarios comunes. AWS CLI los comandos se muestran por motivos de brevedad.

KDC local

Los siguientes comandos crean un clúster con un KDC dedicado del clúster que se ejecuta en el nodo principal. Es necesario realizar una configuración adicional en el clúster. Para obtener más información, consulte Configuración de un clúster de Amazon EMR para conexiones SSH y usuarios de HDFS autenticados en Kerberos.

Crear configuración de seguridad

```
aws emr create-security-configuration --name LocalKDCSecurityConfig \
```

```
--security-configuration '{"AuthenticationConfiguration": \
{"KerberosConfiguration": {"Provider": "ClusterDedicatedKdc",\
"ClusterDedicatedKdcConfiguration": {"TicketLifetimeInHours": 24 }}}}'
```

Crear un clúster

```
aws emr create-cluster --release-label emr-7.10.0 \
--instance-count 3 --instance-type m5.xlarge \
--applications Name=Hadoop Name=Hive --ec2-attributes
 InstanceProfile=EMR_EC2_DefaultRole,KeyName=MyEC2Key \
--service-role EMR_DefaultRole \
--security-configuration LocalKDCSecurityConfig \
--kerberos-attributes Realm=EC2.INTERNAL,KdcAdminPassword=MyPassword
```

KDC dedicado del clúster con una relación de confianza entre ámbitos de Active Directory

Los siguientes comandos crean un clúster con un KDC dedicado del clúster que se ejecuta en el nodo principal con una relación de confianza entre ámbitos con un dominio de Active Directory. Se necesita realizar configuración adicional en el clúster y en Active Directory. Para obtener más información, consulte Tutorial: Configuración de una relación de confianza entre ámbitos con un dominio de Active Directory.

Crear configuración de seguridad

```
aws emr create-security-configuration --name LocalKDCWithADTrustSecurityConfig \
--security-configuration '{"AuthenticationConfiguration": \
{"KerberosConfiguration": {"Provider": "ClusterDedicatedKdc", \
"ClusterDedicatedKdcConfiguration": {"TicketLifetimeInHours": 24, \
"CrossRealmTrustConfiguration": {"Realm":"AD.DOMAIN.COM", \
"Domain":"ad.domain.com", "AdminServer":"ad.domain.com", \
"KdcServer":"ad.domain.com"}}}}}'
```

Crear un clúster

```
aws emr create-cluster --release-label emr-7.10.0 \
--instance-count 3 --instance-type m5.xlarge --applications Name=Hadoop Name=Hive \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,KeyName=MyEC2Key \
--service-role EMR_DefaultRole --security-configuration KDCWithADTrustSecurityConfig \
--kerberos-attributes Realm=EC2.INTERNAL,KdcAdminPassword=MyClusterKDCAdminPassword,\
ADDomainJoinUser=ADUserLogonName,ADDomainJoinPassword=ADUserPassword,\
CrossRealmTrustPrincipalPassword=MatchADTrustPassword
```

KDC externo en un clúster diferente

Los siguientes comandos crean un clúster que hace referencia a un KDC dedicado del clúster en el nodo principal de un clúster diferente para autenticar entidades principales. Es necesario realizar una configuración adicional en el clúster. Para obtener más información, consulte Configuración de un clúster de Amazon EMR para conexiones SSH y usuarios de HDFS autenticados en Kerberos.

Crear configuración de seguridad

```
aws emr create-security-configuration --name ExtKDCOnDifferentCluster \
--security-configuration '{"AuthenticationConfiguration": \
{"KerberosConfiguration": {"Provider": "ExternalKdc", \
"ExternalKdcConfiguration": {"KdcServerType": "Single", \
"AdminServer": "MasterDNSOfKDCMaster:749", \
"KdcServer": "MasterDNSOfKDCMaster:88"}}}}'
```

Crear un clúster

```
aws emr create-cluster --release-label emr-7.10.0 \
--instance-count 3 --instance-type m5.xlarge \
--applications Name=Hadoop Name=Hive \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,KeyName=MyEC2Key \
--service-role EMR_DefaultRole --security-configuration ExtKDCOnDifferentCluster \
--kerberos-attributes Realm=EC2.INTERNAL,KdcAdminPassword=KDCOnMasterPassword
```

KDC externo del clúster con una relación de confianza entre ámbitos de Active Directory

Los siguientes comandos crean un clúster sin KDC. El clúster hace referencia a un KDC dedicado del clúster que se ejecuta en el nodo principal de otro clúster para autenticar entidades principales. Esa KDC posee una relación de confianza entre ámbitos con un controlador de dominio de Active Directory. Es necesario realizar una configuración adicional en el nodo principal con el KDC. Para obtener más información, consulte Tutorial: Configuración de una relación de confianza entre ámbitos con un dominio de Active Directory.

Crear configuración de seguridad

```
aws emr create-security-configuration --name ExtKDCWithADIntegration \
--security-configuration '{"AuthenticationConfiguration": \
{"KerberosConfiguration": {"Provider": "ExternalKdc", \
"ExternalKdcConfiguration": {"KdcServerType": "Single", \
```

```
"AdminServer": "MasterDNSofClusterKDC:749", \
"KdcServer": "MasterDNSofClusterKDC.com:88", \
"AdIntegrationConfiguration": {"AdRealm":"AD.DOMAIN.COM", \
"AdDomain":"ad.domain.com", \
"AdServer":"ad.domain.com"}}}}}'
```

Crear un clúster

```
aws emr create-cluster --release-label emr-7.10.0 \
--instance-count 3 --instance-type m5.xlarge --applications Name=Hadoop Name=Hive \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,KeyName=MyEC2Key \
--service-role EMR_DefaultRole --security-configuration ExtKDCWithADIntegration \
--kerberos-attributes Realm=EC2.INTERNAL,KdcAdminPassword=KDCOnMasterPassword,\
ADDomainJoinUser=MyPrivilegedADUserName,ADDomainJoinPassword=PasswordForADDomainJoinUser
```

Configuración de un clúster de Amazon EMR para conexiones SSH y usuarios de HDFS autenticados en Kerberos

Amazon EMR crea clientes de usuario autenticados por Kerberos para las aplicaciones que se ejecutan en el clúster; por ejemplo, el usuario `hadoop`, el usuario `spark` y otros. También puede añadir usuarios autenticados a los procesos del clúster mediante Kerberos. Luego, los usuarios autenticados pueden conectarse al clúster con sus credenciales de Kerberos y trabajar con aplicaciones. Para que un usuario pueda autenticarse en el clúster, es necesario realizar las siguientes configuraciones:

- En el clúster debe existir una cuenta de Linux que coincida con la entidad principal de Kerberos en el KDC. Amazon EMR lo hace automáticamente en arquitecturas que se integran con Active Directory.
- Debe crear un directorio de usuarios de HDFS en el nodo principal para cada usuario y proporcionar permisos de usuario al directorio.
- Debe configurar el servicio SSH para que GSSAPI esté habilitado en el nodo principal. Además, los usuarios deben tener un cliente SSH con GSSAPI habilitado.

Adición de usuarios de Linux y entidades principales de Kerberos al nodo principal

Si no utiliza Active Directory, debe crear las cuentas de Linux en el nodo principal del clúster y agregar entidades principales para estos usuarios de Linux al KDC. Esto incluye una entidad principal en el KDC para el nodo principal. Además de las entidades principales de usuarios, el KDC que se ejecuta en el nodo principal necesita una entidad principal para el host local.

Cuando la arquitectura incluye integración con Active Directory, los usuarios y entidades principales de Linux en el KDC local, si procede, se crean automáticamente. Puede omitir este paso. Para obtener más información, consulte Relación de confianza entre ámbitos y KDC externo: KDC de clúster en un clúster diferente con una relación de confianza entre ámbitos de Active Directory.

> **Important**
>
> El KDC, junto con la base de datos de entidades principales, se pierde cuando el nodo principal termina porque el nodo principal utiliza almacenamiento efímero. Si crea usuarios para las conexiones SSH, le recomendamos que establezca una relación de confianza entre dominios con un KDC externo configurado para lograr una alta disponibilidad. Como alternativa, si crea usuarios para las conexiones SSH mediante cuentas de Linux, automatice el proceso de creación de cuentas mediante acciones y scripts de arranque para que pueda repetirse al crear un clúster nuevo.

La manera más sencilla de añadir usuarios y entidades principales de KDC es enviar un paso al clúster después de crearlo o al crear el clúster. Como alternativa, puede conectarse al nodo principal mediante un EC2 key pair como `hadoop` usuario predeterminado para ejecutar los comandos. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.

En el siguiente ejemplo, se envía un script bash `configureCluster.sh` a un clúster que ya existe, especificando su ID de clúster. El script se almacena en Amazon S3.

```
aws emr add-steps --cluster-id <j-2AL4XXXXXX5T9> \
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,\
Args=["s3://amzn-s3-demo-bucket/configureCluster.sh"]
```

El siguiente ejemplo muestra el contenido del script `configureCluster.sh`. El script también se encarga de la creación de los directorios de usuario de HDFS y de habilitar GSSAPI para SSH, lo que se explica en las siguientes secciones.

```
#!/bin/bash
#Add a principal to the KDC for the primary node, using the primary node's returned
 host name
sudo kadmin.local -q "ktadd -k /etc/krb5.keytab host/`hostname -f`"
#Declare an associative array of user names and passwords to add
```

```
declare -A arr
arr=([lijuan]=pwd1 [marymajor]=pwd2 [richardroe]=pwd3)
for i in ${!arr[@]}; do
    #Assign plain language variables for clarity
     name=${i}
     password=${arr[${i}]}

     # Create a principal for each user in the primary node and require a new password
 on first logon
     sudo kadmin.local -q "addprinc -pw $password +needchange $name"

     #Add hdfs directory for each user
     hdfs dfs -mkdir /user/$name

     #Change owner of each user's hdfs directory to that user
     hdfs dfs -chown $name:$name /user/$name
done

# Enable GSSAPI authentication for SSH and restart SSH service
sudo sed -i 's/^.*GSSAPIAuthentication.*$/GSSAPIAuthentication yes/' /etc/ssh/
sshd_config
sudo sed -i 's/^.*GSSAPICleanupCredentials.*$/GSSAPICleanupCredentials yes/' /etc/ssh/
sshd_config
sudo systemctl restart sshd
```

Adición de directorios de HDFS de usuarios

Para permitir a los usuarios iniciar sesión en el clúster para ejecutar los trabajos de Hadoop, debe agregar directorios de usuario de HDFS para sus cuentas de Linux y conceder a cada usuario la propiedad de su directorio.

La manera más sencilla de crear directorios de HDFS es enviar un paso al clúster después de crearlo o al crear el clúster. Como alternativa, puede conectarse al nodo principal mediante un EC2 key pair como `hadoop` usuario predeterminado para ejecutar los comandos. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.

En el siguiente ejemplo, se envía un script bash `AddHDFSUsers.sh` a un clúster que ya existe, especificando su ID de clúster. El script se almacena en Amazon S3.

```
aws emr add-steps --cluster-id <j-2AL4XXXXXX5T9> \
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
```

```
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-s3-demo-bucket/AddHDFSUsers.sh"]
```

El siguiente ejemplo muestra el contenido del script `AddHDFSUsers.sh`.

```
#!/bin/bash
# AddHDFSUsers.sh script

# Initialize an array of user names from AD, or Linux users created manually on the
 cluster
ADUSERS=("lijuan" "marymajor" "richardroe" "myusername")

# For each user listed, create an HDFS user directory
# and change ownership to the user

for username in ${ADUSERS[@]}; do
     hdfs dfs -mkdir /user/$username
     hdfs dfs -chown $username:$username /user/$username
done
```

Habilitación de GSSAPI para SSH

Para que los usuarios autenticados por Kerberos se conecten al nodo principal mediante SSH, el servicio SSH debe tener habilitada la autenticación de GSSAPI. Para habilitar GSSAPI, ejecute los siguientes comandos desde la línea de comandos del nodo principal o utilice un paso para ejecutarlo como un script. Después de volver a configurar SSH, reinicie el servicio.

```
sudo sed -i 's/^.*GSSAPIAuthentication.*$/GSSAPIAuthentication yes/' /etc/ssh/
sshd_config
sudo sed -i 's/^.*GSSAPICleanupCredentials.*$/GSSAPICleanupCredentials yes/' /etc/ssh/
sshd_config
sudo systemctl restart sshd
```

Uso de SSH para conectarse a clústeres que utilizan Kerberos con Amazon EMR

En esta sección, se muestran los pasos para que un usuario autenticado por Kerberos se conecte al nodo principal de un clúster de EMR.

Cada equipo que se utiliza para una conexión SSH debe tener instalados un cliente SSH y aplicaciones de cliente de Kerberos. Lo más probable es que los equipos Linux los incluyan de forma

predeterminada. Por ejemplo, OpenSSH está instalado en la mayoría de los sistemas operativos Linux, Unix y macOS. Puede comprobar si tiene un cliente SSH escribiendo ssh en la línea de comandos. Si su equipo no reconoce el comando, instale un cliente SSH para conectarse al nodo principal. El proyecto OpenSSH ofrece una implementación gratuita de toda la suite de herramientas de SSH. Para obtener más información, consulte el sitio web OpenSSH. Los usuarios de Windows pueden utilizar aplicaciones como PuTTY como cliente SSH.

Para obtener más información sobre las conexiones SSH persistentes, use Conectar a un clúster de Amazon EMR.

SSH utiliza GSSAPI para autenticar a los clientes de Kerberos y debe habilitar la autenticación de GSSAPI para el servicio SSH en el nodo principal del clúster. Para obtener más información, consulte Habilitación de GSSAPI para SSH. Los clientes SSH también deben utilizar GSSAPI.

En los siguientes ejemplos, *`MasterPublicDNS`* utilice el valor que aparece para el DNS público maestro en la pestaña Resumen del panel de detalles del clúster, por ejemplo,. *`ec2-11-222-33-44.compute-1.amazonaws.com`*

Requisito previo para krb5.conf (sin Active Directory)

Cuando se utiliza una configuración sin integración con Active Directory, además del cliente SSH y las aplicaciones de cliente de Kerberos, cada equipo cliente debe tener una copia del archivo `/etc/krb5.conf` que coincida con el archivo `/etc/krb5.conf` en el nodo principal del clúster.

Para copiar el archivo krb5.conf

1. Use SSH para conectarse al nodo principal mediante un EC2 key pair y el `hadoop` usuario predeterminado, por ejemplo,. `hadoop@`*`MasterPublicDNS`* Para obtener instrucciones detalladas, consulta Conectar a un clúster de Amazon EMR.
2. Desde el nodo principal, copie el contenido del archivo `/etc/krb5.conf`. Para obtener más información, consulte Conectar a un clúster de Amazon EMR.
3. En cada equipo cliente que se utilice para conectarse al clúster, cree un archivo `/etc/krb5.conf` idéntico a partir de la copia que hizo en el paso anterior.

Uso de kinit y SSH

Cada vez que un usuario se conecta desde un equipo cliente con credenciales de Kerberos, el usuario debe renovar primero los tickets de Kerberos para su usuario en el equipo cliente. Además, se debe configurar el cliente SSH para utilizar la autenticación GSSAPI.

Para utilizar SSH para conectarse a un clúster de EMR que utilice Kerberos

1. Utilice `kinit` para renovar sus tickets de Kerberos, como se muestra en el siguiente ejemplo

```
kinit user1
```

2. Utilice un cliente `ssh` junto con la entidad principal que creó en el KDC dedicado del clúster o el nombre de usuario de Active Directory. Asegúrese de que la autenticación GSSAPI está habilitada, tal y como se muestra en los siguientes ejemplos.

   Ejemplo: usuarios de Linux

   La opción `-K` especifica la autenticación GSSAPI.

```
ssh -K user1@MasterPublicDNS
```

   Ejemplo: usuarios de Windows (PuTTY)

   Asegúrese de que la opción de autenticación GSSAPI de la sesión está habilitada, tal y como se muestra:

Tutorial: Configuración de un KDC dedicado del clúster con Amazon EMR

En este tema, le enseñaremos cómo crear un clúster con un centro de distribución de claves (KDC) dedicado, cómo agregar manualmente cuentas de usuario de Linux a todos los nodos del clúster, cómo agregar entidades principales de Kerberos al KDC en el nodo principal y cómo asegurarse de que los equipos cliente tengan instalado un cliente de Kerberos.

Para obtener más información sobre la compatibilidad de Amazon EMR con Kerberos y KDC, así como enlaces a la documentación de Kerberos del MIT, consulte Uso de Kerberos para la autenticación con Amazon EMR.

Paso 1: creación del clúster que utiliza Kerberos

1. Cree una configuración de seguridad que habilite Kerberos. El siguiente ejemplo muestra un `create-security-configuration` comando que utiliza el AWS CLI que especifica la configuración de seguridad como una estructura JSON en línea. También puede hacer referencia a un archivo guardado de forma local.

```
aws emr create-security-configuration --name MyKerberosConfig \
--security-configuration '{"AuthenticationConfiguration": {"KerberosConfiguration":
{"Provider": "ClusterDedicatedKdc", "ClusterDedicatedKdcConfiguration":
 {"TicketLifetimeInHours": 24}}}}'
```

2. Cree un clúster que haga referencia a la configuración de seguridad, establezca los atributos de Kerberos para el clúster y añada las cuentas de Linux mediante una acción de arranque. El siguiente ejemplo muestra el uso de un comando `create-cluster` en la AWS CLI. El comando hace referencia a la configuración de seguridad que se ha creado anteriormente, `MyKerberosConfig`. También hace referencia a un script sencillo, `createlinuxusers.sh`, como acción de arranque, que debe crear y cargar en Amazon S3 antes de crear el clúster.

```
aws emr create-cluster --name "MyKerberosCluster" \
--release-label emr-7.10.0 \
--instance-type m5.xlarge \
--instance-count 3 \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,KeyName=MyEC2KeyPair \
--service-role EMR_DefaultRole \
--security-configuration MyKerberosConfig \
--applications Name=Hadoop Name=Hive Name=Oozie Name=Hue Name=HCatalog Name=Spark \
--kerberos-attributes Realm=EC2.INTERNAL,\
KdcAdminPassword=MyClusterKDCAdminPwd \
--bootstrap-actions Path=s3://amzn-s3-demo-bucket/createlinuxusers.sh
```

El siguiente código ilustra el contenido del script `createlinuxusers.sh`, que agrega user1, user2 y user3 en cada nodo del clúster. En el siguiente paso, añadirá estos usuarios como entidades principales de KDC.

```
#!/bin/bash
sudo adduser user1
sudo adduser user2
sudo adduser user3
```

Paso 2: adición de entidades principales al KDC, creación directorios de usuarios de HDFS y configuración de SSH

En el KDC que se ejecuta en el nodo principal, se debe agregar una entidad principal para el host local y para cada usuario que se cree en el clúster. También puede crear directorios de HDFS para

cada usuario en caso de que necesitan conectarse al clúster y ejecutar trabajos de Hadoop. Del mismo modo, configurae el servicio SSH para activar la autenticación GSSAPI, que es necesaria para Kerberos. Después de habilitar GSSAPI, reinicie el servicio SSH.

La forma más sencilla de realizar estas tareas es enviar un paso para el clúster. El siguiente ejemplo envía un script bash `configurekdc.sh` al clúster que ha creado en el paso anterior, haciendo referencia a su ID de clúster. El script se almacena en Amazon S3. Como alternativa, puede conectarse al nodo principal mediante un EC2 key pair para ejecutar los comandos o enviar el paso durante la creación del clúster.

```
aws emr add-steps --cluster-id <j-2AL4XXXXXX5T9> --steps
 Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,Jar=s3://
myregion.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-s3-
demo-bucket/configurekdc.sh"]
```

El siguiente código muestra el contenido del script `configurekdc.sh`.

```
#!/bin/bash
#Add a principal to the KDC for the primary node, using the primary node's returned
 host name
sudo kadmin.local -q "ktadd -k /etc/krb5.keytab host/`hostname -f`"
#Declare an associative array of user names and passwords to add
declare -A arr
arr=([user1]=pwd1 [user2]=pwd2 [user3]=pwd3)
for i in ${!arr[@]}; do
    #Assign plain language variables for clarity
     name=${i}
     password=${arr[${i}]}

     # Create principal for sshuser in the primary node and require a new password on
 first logon
     sudo kadmin.local -q "addprinc -pw $password +needchange $name"

     #Add user hdfs directory
     hdfs dfs -mkdir /user/$name

     #Change owner of user's hdfs directory to user
     hdfs dfs -chown $name:$name /user/$name
done

# Enable GSSAPI authentication for SSH and restart SSH service
```

```
sudo sed -i 's/^.*GSSAPIAuthentication.*$/GSSAPIAuthentication yes/' /etc/ssh/
sshd_config
sudo sed -i 's/^.*GSSAPICleanupCredentials.*$/GSSAPICleanupCredentials yes/' /etc/ssh/
sshd_config
sudo systemctl restart sshd
```

Los usuarios que ha añadido ahora deberían ser capaces de conectarse al clúster mediante SSH. Para obtener más información, consulte Uso de SSH para conectarse a clústeres que utilizan Kerberos con Amazon EMR.

Tutorial: Configuración de una relación de confianza entre ámbitos con un dominio de Active Directory

Al crear una relación de confianza entre ámbitos, se permite a las entidades principales (normalmente usuarios) de otro ámbito de Kerberos autenticar a los componentes de la aplicación en el clúster de EMR. El centro de distribución de claves (KDC) dedicado al clúster establece una relación de confianza con otro KDC mediante un principal multidominio que existe en ambos. KDCs El nombre y la contraseña de la entidad principal deben coincidir exactamente.

Una confianza entre dominios requiere que KDCs puedan comunicarse entre sí a través de la red y resolver los nombres de dominio de los demás. A continuación se indican los pasos para establecer una relación de confianza entre dominios con un controlador de dominio de Microsoft AD que se ejecuta como EC2 instancia, junto con un ejemplo de configuración de red que proporciona la conectividad y la resolución de nombres de dominio necesarias. Se acepta cualquier configuración de red que permita el tráfico de red necesario entre KDCs ellos.

Opcionalmente, después de establecer una relación de confianza entre ámbitos con Active Directory utilizando un KDC en un clúster, puede crear otro clúster con una configuración de seguridad diferente para hacer referencia al KDC en el primer clúster como un KDC externo. Para ver un ejemplo de la configuración de seguridad y la configuración del clúster, consulte KDC externo del clúster con una relación de confianza entre ámbitos de Active Directory.

Para obtener más información sobre la compatibilidad de Amazon EMR con Kerberos y KDC, así como enlaces a la documentación de Kerberos del MIT, consulte Uso de Kerberos para la autenticación con Amazon EMR.

> **Important**
>
> Amazon EMR no admite fideicomisos entre dominios con. AWS Directory Service for Microsoft Active Directory

Paso 1: Configuración de la VPC y la subred

Paso 2: lanzamiento e instalación del controlador de dominio de Active Directory

Paso 3: adición de cuentas al dominio para el clúster de EMR

Paso 4: configuración de una relación de confianza entrante en el controlador de dominio de Active Directory

Paso 5: uso de un conjunto de opciones de DHCP para especificar el controlador de dominio de Active Directory como un servidor DNS de la VPC

Paso 6: Lanzar un clúster de EMR que utiliza Kerberos

Paso 7: creación de usuarios de HDFS y configuración de permisos en el clúster para las cuentas de usuario de Active Directory

Paso 1: Configuración de la VPC y la subred

Los pasos siguientes muestran cómo crear una VPC y una subred para que el KDC dedicado del clúster pueda tener acceso al controlador de dominio de Active Directory y resolver su nombre de dominio. En estos pasos, la resolución de nombres de dominio se proporciona haciendo referencia al controlador de dominio de Active Directory como el servidor de nombres de dominio en el conjunto de opciones de DHCP. Para obtener más información, consulte Paso 5: uso de un conjunto de opciones de DHCP para especificar el controlador de dominio de Active Directory como un servidor DNS de la VPC.

El KDC y el controlador de dominio de Active Directory deben ser capaces de resolver sus respectivos nombres de dominio. Esto permite a Amazon EMR incorporar equipos al dominio y configurar automáticamente las cuentas de Linux y los parámetros de SSH correspondientes en las instancias de clúster.

Si Amazon EMR no puede resolver el nombre de dominio, se puede hacer referencia a la relación de confianza mediante la dirección IP del controlador de dominio de Active Directory. Sin embargo,

debe agregar manualmente las cuentas de usuario de Linux, agregar las correspondientes entidades principales al KDC dedicado del clúster y configurar SSH.

Para configurar la VPC y la subred

1. Cree una instancia de Amazon VPC con una única subred pública. Para obtener más información, consulte Paso 1: creación de la VPC en la Guía de introducción a Amazon VPC.

   > **Important**
   >
   > Cuando utilice un controlador de dominio de Microsoft Active Directory, elija un bloque CIDR para el clúster de EMR de modo que IPv4 todas las direcciones tengan menos de nueve caracteres (por ejemplo, 10.0.0.0/16). Esto se debe a que los nombres DNS de los equipos del clúster se utilizan cuando los equipos se unen al directorio de Active Directory. AWS asigna nombres de host DNS en función de la IPv4 dirección, de manera que las direcciones IP más largas pueden dar como resultado nombres DNS de más de 15 caracteres. Active Directory tiene un límite de 15 caracteres para registrar los nombres de los equipos que se incorporan, y trunca los nombres que son más largos, lo que puede provocar errores impredecibles.

2. Elimine el conjunto de opciones de DHCP predeterminado asignado a la VPC. Para obtener más información, consulte Cambiar una VPC para que no utilice ninguna opción de DHCP. Posteriormente, puede añadir uno nuevo que especifique el controlador de dominio de Active Directory como servidor DNS.
3. Confirme que se ha activado el soporte de DNS para la VPC, es decir, que se han activado los nombres de host DNS y la resolución de DNS. De forma predeterminada, están habilitadas. Para obtener más información, consulte Actualización de la compatibilidad de DNS para su VPC.
4. Confirme que la VPC tiene un asociada una gateway de Internet, que es la opción predeterminada. Para obtener más información, consulte Creación y asociación de una gateway de Internet.

   > **Note**
   >
   > En este ejemplo, se utiliza una gateway de Internet porque se está estableciendo un nuevo controlador de dominio para la VPC. Puede que no sea necesaria una gateway de Internet para su aplicación. El único requisito es que el KDC dedicado del clúster pueda tener acceso al controlador de dominio de Active Directory.

5. Cree una tabla de ruteo personalizada, añada una ruta que se dirija a la gateway de Internet y, a continuación, asóciela a la subred. Para obtener más información, consulte Creación de una tabla de enrutamiento personalizada.
6. Al lanzar la EC2 instancia para el controlador de dominio, debe tener una IPv4 dirección pública estática para que pueda conectarse a ella mediante RDP. La forma más sencilla de hacerlo es configurar la subred para que asigne automáticamente direcciones públicas IPv4 . Este no es el valor predeterminado cuando se crea una subred. Para obtener más información, consulte Modificación del atributo de IPv4 direccionamiento público de la subred. Si lo prefiere, puede asignar la dirección al lanzar la instancia. Para obtener más información, consulta Cómo asignar una IPv4 dirección pública durante el lanzamiento de una instancia.
7. Cuando termine, anote la VPC y la subred. IDs Los utilizará posteriormente al lanzar el controlador de dominio de Active Directory y el clúster.

Paso 2: lanzamiento e instalación del controlador de dominio de Active Directory

1. Lance una EC2 instancia basada en la AMI básica de Microsoft Windows Server 2016. Le recomendamos un tipo de instancia m4.xlarge o mejor. Para obtener más información, consulta Cómo lanzar una AWS Marketplace instancia en la Guía del EC2 usuario de Amazon.
2. Anote el ID de grupo del grupo de seguridad asociado a la EC2 instancia. Lo necesitará para Paso 6: Lanzar un clúster de EMR que utiliza Kerberos. Usamos`sg-012xrlmdomain345`. También puede especificar distintos grupos de seguridad para el clúster de EMR y esta instancia que permita el tráfico entre ellos. Para obtener más información, consulta los grupos EC2 de seguridad de Amazon para instancias de Linux en la Guía del EC2 usuario de Amazon.
3. Conéctate a la EC2 instancia mediante RDP. Para obtener más información, consulta Cómo conectarse a tu instancia de Windows en la Guía del EC2 usuario de Amazon.
4. Inicie Administrador del servidor para instalar y configurar el rol de los servicios de dominio de Active Directory en el servidor. Promocione el servidor a controlador de dominio y asígnele un nombre de dominio (el ejemplo que utilizamos aquí es `ad.domain.com`). Anote el nombre de dominio porque lo necesitará más adelante al crear el clúster y la configuración de seguridad de EMR. Si es la primera vez que configura Active Directory, puede seguir las instrucciones de How to Set Up Active Directory (AD) in Windows Server 2016.

   La instancia se reiniciará cuando termine.

Paso 3: adición de cuentas al dominio para el clúster de EMR

Establezca una conexión RDP con el controlador de dominio de Active Directory para crear cuentas de usuario en Usuarios y equipos de Active Directory para cada usuario del clúster. Para obtener más información, consulte Create a User Account in Active Directory Users and Computers en el sitio Microsoft Learn. Anote el User logon name (Nombre de inicio de sesión de usuario) de cada usuario. Necesitas estas versiones posteriores al configurar el clúster.

Cree también una cuenta con privilegios suficientes para incorporar ordenadores al dominio. Tiene que especificar esta cuenta al crear un clúster. Amazon EMR la utiliza para incorporar las instancias del clúster al dominio Solo debe especificar esta cuenta y su contraseña Paso 6: Lanzar un clúster de EMR que utiliza Kerberos. Para delegar los privilegios de incorporación de equipos a la cuenta, le recomendamos que cree un grupo con privilegios de incorporación y, a continuación, asigne el usuario al grupo. Para obtener instrucciones, consulte Delegación de privilegios de vinculación a directorios en la Guía de administración de AWS Directory Service .

Paso 4: configuración de una relación de confianza entrante en el controlador de dominio de Active Directory

Los comandos del ejemplo siguiente crean una relación de confianza en Active Directory, que es una confianza de ámbito, entrante, unidireccional y no transitiva con el KDC dedicado del clúster. El ejemplo que utilizamos para el ámbito del clúster es *`EC2.INTERNAL`*. Sustituya el *`KDC-FQDN`* nombre de DNS público que aparece en la lista para el nodo principal de Amazon EMR que aloja el KDC. El parámetro `passwordt` especifica la cross-realm principal password (contraseña de la entidad principal de confianza entre ámbitos), que se especifica junto con el realm (ámbito) del clúster al crear un clúster. El nombre del ámbito se deriva del nombre de dominio predeterminado en `us-east-1` para el clúster. El `Domain` es el dominio de Active Directory en el que va a crear la confianza, que es en minúsculas por convención. El ejemplo utiliza *`ad.domain.com`*

Abra el símbolo del sistema de Windows con privilegios de administrador y escriba los siguientes comandos para crear la relación de confianza en el controlador de dominio de Active Directory:

```
C:\Users\Administrator> ksetup /addkdc EC2.INTERNAL KDC-FQDN
C:\Users\Administrator> netdom trust EC2.INTERNAL /Domain:ad.domain.com /add /realm /
passwordt:MyVeryStrongPassword
C:\Users\Administrator> ksetup /SetEncTypeAttr EC2.INTERNAL AES256-CTS-HMAC-SHA1-96
```

Paso 5: uso de un conjunto de opciones de DHCP para especificar el controlador de dominio de Active Directory como un servidor DNS de la VPC

Ahora que está configurado el controlador de dominio de Active Directory, debe configurar la VPC para utilizarla como servidor de nombres de dominio para la resolución de nombres en la VPC. Para ello, asocie un conjunto de opciones de DHCP. En Nombre de dominio, especifique el nombre de dominio del clúster; por ejemplo, `ec2.internal` si el clúster se encuentra en la región us-east-1 o *`region`*`.compute.internal` para las demás regiones. En el caso de los servidores de nombres de dominio, debe especificar la dirección IP del controlador de dominio de Active Directory (al que se debe poder acceder desde el clúster) como primera entrada, seguida del AmazonProvidedDNS (por ejemplo ***`xx.xx.xx.xx`***, AmazonProvided DNS). Para obtener más información, consulte Cambio de los conjuntos de opciones de DHCP.

Paso 6: Lanzar un clúster de EMR que utiliza Kerberos

1. En Amazon EMR, cree una configuración de seguridad que especifique el controlador de dominio de Active Directory que ha creado en los pasos anteriores. A continuación se muestra un ejemplo. Sustituya el dominio, *`ad.domain.com`*por el nombre del dominio especificado en el Paso 2: lanzamiento e instalación del controlador de dominio de Active Directory.

```
aws emr create-security-configuration --name MyKerberosConfig \
--security-configuration '{
  "AuthenticationConfiguration": {
    "KerberosConfiguration": {
      "Provider": "ClusterDedicatedKdc",
      "ClusterDedicatedKdcConfiguration": {
        "TicketLifetimeInHours": 24,
        "CrossRealmTrustConfiguration": {
          "Realm": "AD.DOMAIN.COM",
          "Domain": "ad.domain.com",
          "AdminServer": "ad.domain.com",
          "KdcServer": "ad.domain.com"
        }
      }
    }
  }
}'
```

2. Cree el clúster con los siguientes atributos:
   - Utilice la opción `--security-configuration` para especificar la configuración de seguridad que ha creado. Usamos *`MyKerberosConfig`* en el ejemplo.

- Utilice la propiedad `SubnetId` de `--ec2-attributes option` para especificar la subred que ha creado en Paso 1: Configuración de la VPC y la subred. Usamos *`step1-subnet`* en el ejemplo.
- Utilice `AdditionalMasterSecurityGroups` y `AdditionalSlaveSecurityGroups` de la opción `--ec2-attributes` para especificar que el grupo de seguridad asociado al controlador de dominio de AD desde Paso 2: lanzamiento e instalación del controlador de dominio de Active Directory está asociado al nodo principal del clúster, así como a los nodos secundarios y de tareas. Usamos *`sg-012xrlmdomain345`* en el ejemplo.

Utilice `--kerberos-attributes` para especificar los siguientes atributos de Kerberos específicos del clúster:

- El ámbito para el clúster que especificó al configurar el controlador de dominio de Active Directory.
- La contraseña de la entidad principal de confianza entre ámbitos que especificó como `passwordt` en el Paso 4: configuración de una relación de confianza entrante en el controlador de dominio de Active Directory.
- Una `KdcAdminPassword`, que se puede utilizar para administrar el KDC dedicado del clúster.
- El nombre de inicio de sesión de usuario y la contraseña de la cuenta de Active Directory con privilegios para incorporar equipos que ha creado en el Paso 3: adición de cuentas al dominio para el clúster de EMR.

El siguiente ejemplo lanza un clúster que utiliza Kerberos.

```
aws emr create-cluster --name "MyKerberosCluster" \
--release-label emr-5.10.0 \
--instance-type m5.xlarge \
--instance-count 3 \
--ec2-attributes InstanceProfile=EMR_EC2_DefaultRole,KeyName=MyEC2KeyPair,\
SubnetId=step1-subnet, AdditionalMasterSecurityGroups=sg-012xrlmdomain345,
AdditionalSlaveSecurityGroups=sg-012xrlmdomain345\
--service-role EMR_DefaultRole \
--security-configuration MyKerberosConfig \
--applications Name=Hadoop Name=Hive Name=Oozie Name=Hue Name=HCatalog Name=Spark \
--kerberos-attributes Realm=EC2.INTERNAL,\
KdcAdminPassword=MyClusterKDCAdminPwd,\
ADDomainJoinUser=ADUserLogonName,ADDomainJoinPassword=ADUserPassword,\
CrossRealmTrustPrincipalPassword=MatchADTrustPwd
```

Paso 7: creación de usuarios de HDFS y configuración de permisos en el clúster para las cuentas de usuario de Active Directory

Cuando se configura una relación de confianza con Active Directory, Amazon EMR crea usuarios de Linux en el clúster para cada cuenta de usuario de Active Directory. Por ejemplo, el nombre de inicio de sesión de usuario `LiJuan` en Active Directory se corresponde con la cuenta de Linux de `lijuan`. Los nombres de usuario de Active Directory pueden contener letras mayúsculas, pero Linux no distingue entre mayúsculas y minúsculas para dichos nombres.

Para permitir a los usuarios iniciar sesión en el clúster para ejecutar los trabajos de Hadoop, debe agregar directorios de usuario de HDFS para sus cuentas de Linux y conceder a cada usuario la propiedad de su directorio. Para ello, le recomendamos que ejecute un script almacenado en Amazon S3 como un paso de clúster. De forma alternativa, puede ejecutar los comandos en el script siguiente desde la línea de comandos en el nodo principal. Use el par de EC2 claves que especificó al crear el clúster para conectarse al nodo principal a través de SSH como usuario de Hadoop. Para obtener más información, consulte Utilice un par de EC2 claves para las credenciales SSH de Amazon EMR.

Ejecute el siguiente comando para añadir un paso al clúster que ejecuta un script,. *`AddHDFSUsers.sh`*

```
aws emr add-steps --cluster-id <j-2AL4XXXXXX5T9> \
--steps Type=CUSTOM_JAR,Name=CustomJAR,ActionOnFailure=CONTINUE,\
Jar=s3://region.elasticmapreduce/libs/script-runner/script-runner.jar,Args=["s3://amzn-s3-demo-bucket/AddHDFSUsers.sh"]
```

El contenido del archivo *`AddHDFSUsers.sh`* es el siguiente.

```
#!/bin/bash
# AddHDFSUsers.sh script

# Initialize an array of user names from AD or Linux users and KDC principals created
 manually on the cluster
ADUSERS=("lijuan" "marymajor" "richardroe" "myusername")

# For each user listed, create an HDFS user directory
# and change ownership to the user

for username in ${ADUSERS[@]}; do
     hdfs dfs -mkdir /user/$username
     hdfs dfs -chown $username:$username /user/$username
```

```
done
```

Grupos de Active Directory asignados a grupos de Hadoop

Amazon EMR utiliza System Security Services Daemon (SSD) para asignar grupos de Active Directory a grupos de Hadoop. Para confirmar las asignaciones de grupos, después de iniciar sesión en el nodo principal tal y como se describe en Uso de SSH para conectarse a clústeres que utilizan Kerberos con Amazon EMR, puede utilizar el comando `hdfs groups` para confirmar que los grupos de Active Directory a los que pertenece su cuenta de Active Directory se han asignado a los grupos de Hadoop del usuario de Hadoop correspondiente en el clúster. También puede comprobar los mapeos de grupos de otros usuarios especificando uno o varios nombres de usuario con el comando, por ejemplo `hdfs groups lijuan`. Para obtener más información, consulte groups en la Apache HDFS Commands Guide.

Uso de servidores de Active Directory o LDAP para la autenticación con Amazon EMR

Con las versiones 6.12.0 y posteriores de Amazon EMR, puede utilizar el protocolo LDAP sobre SSL (LDAPS) para lanzar un clúster que se integre de forma nativa con su servidor de identidad corporativa. El LDAP (protocolo ligero de acceso a directorios) es un protocolo de aplicaciones abierto e independiente del proveedor que accede a los datos y los mantiene. El LDAP se utiliza habitualmente para la autenticación de usuarios en servidores de identidad corporativa alojados en aplicaciones como Active Directory (AD) y OpenLDAP. Con esta integración nativa, puede usar su servidor LDAP para autenticar a los usuarios en Amazon EMR.

Entre los aspectos destacados de la integración del LDAP de Amazon EMR se encuentran los siguientes:

- Amazon EMR configura las aplicaciones compatibles para que se autentiquen con la autenticación LDAP en su nombre.
- Amazon EMR configura y mantiene la seguridad de las aplicaciones compatibles con el protocolo Kerberos. No es necesario introducir ningún comando o script.
- Se obtiene un control de acceso detallado (FGAC) mediante la autorización de Apache Ranger para las bases de datos y tablas del metaalmacén de Hive. Para obtener más información, consulte Integración de Amazon EMR con Apache Ranger.
- Cuando necesita credenciales del LDAP para acceder a un clúster, obtiene un control de acceso detallado (FGAC) sobre quién puede acceder a sus clústeres de EMR a través de SSH.

En las siguientes páginas se proporciona información general conceptual, los requisitos previos y los pasos para lanzar un clúster de EMR con la integración del LDAP de Amazon EMR.

Temas

- Descripción general del LDAP con Amazon EMR
- Componentes del LDAP para Amazon EMR
- Compatibilidad de aplicaciones y consideraciones que hay que tener en cuenta sobre el LDAP para Amazon EMR
- Configuración y lanzamiento de un clúster de EMR con el LDAP
- Ejemplos de uso del LDAP con Amazon EMR

Descripción general del LDAP con Amazon EMR

El protocolo ligero de acceso a directorios (LDAP) es un protocolo de software que los administradores de red utilizan para administrar y controlar el acceso a los datos mediante la autenticación de los usuarios dentro de la red de una empresa. El protocolo LDAP almacena la información en una estructura jerárquica de directorios en árbol. Para obtener más información, consulte Basic LDAP Concepts en LDAP.com.

Dentro de la red de una empresa, muchas aplicaciones pueden utilizar el protocolo LDAP para autenticar a los usuarios. Con la integración del LDAP de Amazon EMR, los clústeres de EMR pueden utilizar de forma nativa el mismo protocolo LDAP con una configuración de seguridad adicional.

Amazon EMR admite dos implementaciones principales del protocolo LDAP: Active Directory y OpenLDAP. Si bien son posibles otras implementaciones, la mayoría se ajusta a los mismos protocolos de autenticación que Active Directory u OpenLDAP.

Active Directory (AD)

Active Directory (AD) es un servicio de directorios de Microsoft para redes de dominio de Windows. AD está incluido en la mayoría de los sistemas operativos Windows Server y puede comunicarse con los clientes a través de los protocolos LDAP y LDAPS. Amazon EMR intenta vincular un usuario a su instancia de AD con el nombre principal del usuario (UPN) como nombre distintivo y la contraseña para realizar la autenticación. El UPN utiliza el formato estándar `username@domain_name`.

OpenLDAP

OpenLDAP es una implementación gratuita y de código abierto del protocolo LDAP. Para la autenticación, Amazon EMR intenta vincular un usuario a su instancia de OpenLDAP con el nombre de dominio completo (FQDN) como nombre distintivo y la contraseña. El FQDN utiliza el formato estándar `username_attribute=username,LDAP_user_search_base`. Por lo general, el valor `username_attribute` es `uid` y el valor `LDAP_user_search_base` contiene los atributos del árbol que conduce al usuario. Por ejemplo, `ou=People,dc=example,dc=com`.

Otras implementaciones gratuitas y de código abierto del protocolo LDAP suelen seguir un FQDN similar al de OpenLDAP para los nombres distintivos de sus usuarios.

Componentes del LDAP para Amazon EMR

Puede usar su servidor LDAP para autenticarse con Amazon EMR y cualquier aplicación que el usuario utilice directamente en el clúster de EMR a través de los siguientes componentes.

Agente secreto

El agente secreto es un proceso que se ejecuta en el clúster que autentica todas las solicitudes de los usuarios. El agente secreto crea el enlace de usuario a su servidor LDAP en nombre de las aplicaciones compatibles en el clúster de EMR. El agente secreto se ejecuta como el usuario `emrsecretagent` y escribe registros en el directorio `/emr/secretagent/log`. Estos registros proporcionan detalles sobre el estado de la solicitud de autenticación de cada usuario y cualquier error que pueda surgir durante la autenticación del usuario.

System Security Services Daemon (SSSD)

SSSD es un daemon que se ejecuta en cada nodo de un clúster de EMR habilitado para el LDAP. SSSD crea y administra un usuario de UNIX para sincronizar su identidad corporativa remota con cada nodo. Las aplicaciones basadas en YARN, como Hive y Spark, requieren que haya un usuario local de UNIX en cada nodo que ejecute una consulta para un usuario.

Compatibilidad de aplicaciones y consideraciones que hay que tener en cuenta sobre el LDAP para Amazon EMR

En este tema se enumeran las aplicaciones compatibles, las características compatibles y las características no compatibles.

Aplicaciones compatibles con el LDAP para Amazon EMR

> **Important**
>
> Las aplicaciones que aparecen en esta página son las únicas aplicaciones que Amazon EMR admite para el LDAP. Para garantizar la seguridad del clúster, solo puede incluir aplicaciones compatibles con el LDAP al crear un clúster de EMR con el LDAP habilitado. Si intenta instalar otras aplicaciones no compatibles, Amazon EMR rechazará su solicitud de un nuevo clúster.

Las versiones 6.12 y posteriores de Amazon EMR admiten la integración del LDAP con las siguientes aplicaciones:

- Apache Livy
- Desde Apache Hive hasta HiveServer 2 () HS2
- Trino
- Presto
- Hue

También puede instalar las siguientes aplicaciones en un clúster de EMR y configurarlas para que se ajusten a sus necesidades de seguridad:

- Apache Spark
- Apache Hadoop

Características compatibles con el LDAP para Amazon EMR

Puede utilizar las siguientes características de Amazon EMR con la integración del LDAP:

> **Note**
>
> Para mantener seguras las credenciales del LDAP, debe utilizar el cifrado en tránsito para proteger el flujo de datos dentro y fuera del clúster. Para obtener más información sobre el cifrado en tránsito, consulte Cifrar datos en reposo y en tránsito con Amazon EMR.

- Cifrado en reposo y en tránsito (obligatorio)
- Grupos de instancias, flotas de instancias e instancias de spot
- Reconfiguración de aplicaciones en un clúster en ejecución
- Cifrado del servidor (SSE) de EMRFS

Características no admitidas

Tenga en cuenta las siguientes limitaciones cuando utilice la integración del LDAP de Amazon EMR:

- Amazon EMR deshabilita los pasos para los clústeres con el LDAP habilitado.
- Amazon EMR no admite funciones e AWS Lake Formation integraciones de tiempo de ejecución para clústeres con LDAP habilitado.
- Amazon EMR no admite el LDAP con StartTLS.
- Amazon EMR no admite el modo de alta disponibilidad (clústeres con varios nodos principales) en los clústeres con el LDAP habilitado.
- No puede vincular de forma rotativa las credenciales ni los certificados de los clústeres con el LDAP habilitado. Si alguno de esos campos se ha rotado, le recomendamos que inicie un clúster nuevo con las credenciales o certificados de vinculación actualizados.
- Debe utilizar bases de búsqueda exactas con LDAP. La base de búsqueda de usuarios y grupos de LDAP no admite los filtros de búsqueda de LDAP.

Configuración y lanzamiento de un clúster de EMR con el LDAP

En esta sección, se explica cómo configurar Amazon EMR para su uso con la autenticación LDAP.

Temas

- Añadir AWS Secrets Manager permisos al rol de instancia de Amazon EMR
- Creación de la configuración de seguridad de Amazon EMR para la integración con el LDAP
- Lanzamiento de un clúster de EMR que se autentique con LDAP

Añadir AWS Secrets Manager permisos al rol de instancia de Amazon EMR

Amazon EMR utiliza un rol de servicio de IAM para realizar acciones en su nombre para aprovisionar y administrar clústeres. La función de servicio para EC2 las instancias de clúster, también

denominada perfil de EC2 instancia de Amazon EMR, es un tipo especial de función de servicio que Amazon EMR asigna a todas las EC2 instancias de un clúster en el momento del lanzamiento.

Para definir los permisos para que un clúster de EMR interactúe con los datos de Amazon S3 y otros AWS servicios, defina un perfil de EC2 instancia de Amazon personalizado en lugar del que tenía `EMR_EC2_DefaultRole` al lanzar el clúster. Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia) y Personalización de roles de IAM con Amazon EMR.

Añada las siguientes instrucciones al perfil de EC2 instancia predeterminado para permitir que Amazon EMR etiquete las sesiones y acceda al AWS Secrets Manager que almacena los certificados LDAP.

```
    {
      "Sid": "AllowAssumeOfRolesAndTagging",
      "Effect": "Allow",
      "Action": ["sts:TagSession", "sts:AssumeRole"],
      "Resource": [
        "arn:aws:iam::111122223333:role/LDAP_DATA_ACCESS_ROLE_NAME",
        "arn:aws:iam::111122223333:role/LDAP_USER_ACCESS_ROLE_NAME"
      ]
    },
    {
        "Sid": "AllowSecretsRetrieval",
        "Effect": "Allow",
        "Action": "secretsmanager:GetSecretValue",
        "Resource": [
            "arn:aws:secretsmanager:us-east-1:111122223333:secret:LDAP_SECRET_NAME*",
            "arn:aws:secretsmanager:us-
east-1:111122223333:secret:ADMIN_LDAP_SECRET_NAME*"
        ]
    }
```

> **Note**
>
> Las solicitudes de clúster fallarán si olvida agregar el carácter comodín * al final del nombre del secreto al configurar los permisos de Secrets Manager. El comodín representa las versiones secretas.
> También debes limitar el alcance de la AWS Secrets Manager política a solo los certificados que tu clúster necesita para aprovisionar instancias.

Creación de la configuración de seguridad de Amazon EMR para la integración con el LDAP

Antes de lanzar un clúster de EMR con integración del LDAP, siga los pasos de Cree una configuración de seguridad con la consola Amazon EMR o con AWS CLI para crear una configuración de seguridad de Amazon EMR para el clúster. Complete las siguientes configuraciones en el bloque `LDAPConfiguration` en `AuthenticationConfiguration` o en los campos correspondientes de la sección Configuraciones de seguridad de la consola de Amazon EMR:

`EnableLDAPAuthentication`

Opción de consola: Protocolo de autenticación: LDAP

Para usar la integración del LDAP, establezca esta opción en `true` o selecciónela como protocolo de autenticación al crear un clúster en la consola. De forma predeterminada, `EnableLDAPAuthentication` es `true` cuando crea una configuración de seguridad en la consola de Amazon EMR.

`LDAPServerURL`

Opción de consola: ubicación del servidor LDAP

La ubicación del servidor LDAP, incluido el prefijo: `ldaps://`*`location_of_server`*.

`BindCertificateARN`

Opción de consola: Certificado SSL del LDAP

El AWS Secrets Manager ARN que contiene el certificado para firmar el certificado SSL que utiliza el servidor LDAP. Si su servidor LDAP está firmado por una autoridad de certificación (CA) pública, puede proporcionar un AWS Secrets Manager ARN con un archivo en blanco. Para obtener más información sobre cómo almacenar el certificado en Secrets Manager, consulte Almacenamiento de certificados TLS en AWS Secrets Manager.

`BindCredentialsARN`

Opción de consola: credenciales de vinculación del servidor LDAP

Un AWS Secrets Manager ARN que contiene las credenciales de enlace de usuario administrador de LDAP. Las credenciales se almacenan como un objeto JSON. Solo hay un par clave-valor en este secreto; la clave del par es el nombre de usuario y el valor es la contraseña. Por ejemplo, `{"uid=admin,cn=People,dc=example,dc=com": "AdminPassword1"}`. Este campo es opcional, a menos que habilite el inicio de sesión SSH para su clúster de EMR. En muchas

configuraciones, las instancias de Active Directory requieren credenciales de vinculación para permitir que SSSD sincronice los usuarios.

`LDAPAccessFilter`

Opción de consola: filtro de acceso del LDAP

Especifica el subconjunto de objetos del servidor LDAP que pueden autenticarse. Por ejemplo, si desea conceder acceso a todos los usuarios con la clase de objeto `posixAccount` de su servidor LDAP, defina el filtro de acceso como `(objectClass=posixAccount)`.

`LDAPUserSearchBase`

Opción de consola: base de búsqueda de usuarios de LDAP

La base de búsqueda a la que pertenecen sus usuarios en su servidor LDAP. Por ejemplo, `cn=People,dc=example,dc=com`.

`LDAPGroupSearchBase`

Opción de consola: base de búsqueda de grupos LDAP

La base de búsqueda a la que pertenecen sus grupos dentro de su servidor LDAP. Por ejemplo, `cn=Groups,dc=example,dc=com`.

`EnableSSHLogin`

Opción de consola: inicio de sesión SSH

Especifica si se permite o no la autenticación por contraseña con credenciales del LDAP. No le recomendamos habilitar esta opción. Los pares de claves son una ruta más segura para permitir el acceso a los clústeres de EMR. Este campo es opcional y de forma predeterminada es `false`.

`LDAPServerType`

Opción de consola: tipo de servidor LDAP

Especifica el tipo de servidor LDAP al que se conecta Amazon EMR. Las opciones compatibles son Active Directory y OpenLDAP. Es posible que otros tipos de servidores LDAP funcionen, pero Amazon EMR no admite oficialmente otros tipos de servidores. Para obtener más información, consulte Componentes del LDAP para Amazon EMR.

`ActiveDirectoryConfigurations`

Un subbloque obligatorio para las configuraciones de seguridad que utilizan el tipo de servidor Active Directory.

`ADDomain`

Opción de consola: dominio de Active Directory

El nombre de dominio utilizado para crear el nombre principal de usuario (UPN) para la autenticación del usuario con configuraciones de seguridad que utilizan el tipo de servidor Active Directory.

Aspectos que hay que tener en cuenta para realizar las configuraciones de seguridad con el LDAP y Amazon EMR

- Para crear una configuración de seguridad con la integración del LDAP de Amazon EMR, debe utilizar el cifrado en tránsito. Para obtener información sobre el cifrado en tránsito, consulte Cifrar datos en reposo y en tránsito con Amazon EMR.
- No puede definir la configuración de Kerberos en la misma configuración de seguridad. Amazon EMR aprovisiona un KDC dedicado al KDC de forma automática y administra la contraseña de administrador de este KDC. Los usuarios no pueden acceder a esta contraseña de administrador.
- No puede definir las funciones de tiempo de ejecución de IAM ni AWS Lake Formation en la misma configuración de seguridad.
- La `LDAPServerURL` debe tener el protocolo `ldaps://` en su valor.
- El `LDAPAccessFilter` no puede estar vacío.

Uso del LDAP con la integración de Apache Ranger para Amazon EMR

Con la integración del LDAP para Amazon EMR, puede ampliar su integración con Apache Ranger. Cuando extraiga sus usuarios del LDAP en Ranger, podrá asociarlos a un servidor de políticas de Apache Ranger para integrarlos con Amazon EMR y otras aplicaciones. Para ello, defina el campo `RangerConfiguration` de la configuración de seguridad que `AuthorizationConfiguration` va a utilizar con el clúster del LDAP. Para obtener más información acerca de cómo definir la configuración de seguridad, consulte Creación de la configuración de seguridad de EMR.

Cuando utilice el LDAP con Amazon EMR, no tendrá que proporcionar una `KerberosConfiguration` con la integración de Amazon EMR para Apache Ranger.

Lanzamiento de un clúster de EMR que se autentique con LDAP

Siga estos pasos para lanzar un clúster de EMR con el LDAP o Active Directory:

1. Configure su entorno:
   - Asegúrese de que los nodos del clúster de EMR puedan comunicarse con Amazon S3 y. AWS Secrets Manager Para obtener más información sobre cómo modificar el rol de su perfil de EC2 instancia para comunicarse con estos servicios, consulteAñadir AWS Secrets Manager permisos al rol de instancia de Amazon EMR.
   - Si planea ejecutar su clúster de EMR en una subred privada, debe usar puntos de enlace de AWS PrivateLink Amazon VPC o usar la traducción de direcciones de red (NAT) para configurar la VPC de manera que se comunique con S3 y Secrets Manager. Para obtener más información, consulte AWS PrivateLink y puntos de conexión de VPC e Instancias NAT en la Guía de introducción a Amazon VPC.
   - Asegúrese de que haya conectividad de red entre el clúster de EMR y el servidor LDAP. Los clústeres de EMR deben acceder a su servidor LDAP a través de la red. Los nodos principal, central y de tareas del clúster se comunican con el servidor LDAP para sincronizar los datos de los usuarios. Si su servidor LDAP se ejecuta en Amazon EC2, actualice el grupo de EC2 seguridad para que acepte el tráfico del clúster EMR. Para obtener más información, consulte Añadir AWS Secrets Manager permisos al rol de instancia de Amazon EMR.
2. Cree una configuración de seguridad de Amazon EMR para la integración del LDAP. Para obtener más información, consulte Creación de la configuración de seguridad de Amazon EMR para la integración con el LDAP.
3. Una vez completada la configuración, siga los pasos de Lanzar un clúster de Amazon EMR para lanzar el clúster con las siguientes configuraciones:
   - Seleccione Amazon EMR versión 6.12 o superior. Le recomendamos utilizar la última versión de Amazon EMR.
   - Especifique o seleccione únicamente las aplicaciones de su clúster que admitan el LDAP. Para obtener una lista de las aplicaciones que admitan el LDAP con Amazon EMR, consulte Compatibilidad de aplicaciones y consideraciones que hay que tener en cuenta sobre el LDAP para Amazon EMR.
   - Aplique la configuración de seguridad que creó en el paso anterior.

Ejemplos de uso del LDAP con Amazon EMR

Una vez que aprovisione un clúster de EMR que utilice la integración del LDAP, podrá proporcionar sus credenciales de LDAP a cualquier aplicación compatible mediante su mecanismo de autenticación de nombre de usuario y contraseña integrado. Esta página muestra algunos ejemplos.

Uso de la autenticación LDAP con Apache Hive

Example : Apache Hive

El siguiente comando de ejemplo inicia una sesión de Apache Hive a través HiveServer de 2 y Beeline:

```
beeline -u "jdbc:hive2:///$HOSTNAME:10000/default;ssl=true;sslTrustStore=
$TRUSTSTORE_PATH;trustStorePassword=$TRUSTSTORE_PASS"  -n LDAP_USERNAME -
p LDAP_PASSWORD
```

Uso de la autenticación LDAP con Apache Livy

Example : Apache Livy

El siguiente comando de ejemplo inicia una sesión de Livy a través de cURL. Sustituya *ENCODED-KEYPAIR* por una cadena cifrada en Base64 para `username:password`.

```
curl -X POST --data '{"proxyUser":"LDAP_USERNAME","kind": "pyspark"}' -H "Content-Type:
 application/json" -H "Authorization: Basic ENCODED-KEYPAIR" DNS_OF_PRIMARY_NODE:8998/
sessions
```

Uso de la autenticación LDAP con Presto

Example : Presto

El siguiente comando de ejemplo inicia una sesión de Presto a través de la CLI de Presto:

```
presto-cli --user "LDAP_USERNAME" --password --catalog hive
```

Tras ejecutar este comando, ingrese la contraseña de LDAP cuando se solicite.

Uso de la autenticación LDAP con Trino

Example : Trino

El siguiente comando de ejemplo inicia una sesión de Trino a través de la CLI de Trino:

```
trino-cli --user "LDAP_USERNAME" --password --catalog hive
```

Tras ejecutar este comando, ingrese la contraseña de LDAP cuando se solicite.

Uso de la autenticación LDAP con Hue

Puede acceder a la interfaz de usuario de Hue a través de un túnel SSH que cree en el clúster, o bien puede configurar un servidor proxy para que transmita públicamente la conexión a Hue. Como Hue no se ejecuta en modo HTTPS de forma predeterminada, le recomendamos que utilice una capa de cifrado adicional para garantizar que la comunicación entre los clientes y la interfaz de usuario de Hue esté cifrada con HTTPS. Esto reduce la posibilidad de que exponga accidentalmente las credenciales de usuario en texto sin formato.

Para usar la interfaz de usuario de Hue, abra la interfaz de usuario de Hue en su navegador e ingrese la contraseña y su nombre de usuario del LDAP para iniciar sesión. Si las credenciales son correctas, Hue inicia sesión y usa su identidad para que pueda autenticarse en todas las aplicaciones compatibles.

Uso de SSH para la autenticación por contraseña y tickets de Kerberos para otras aplicaciones

> **Important**
>
> No le recomendamos autenticarse por contraseña para conectarse vía SSH a un clúster de EMR.

Puede usar sus credenciales del LDAP para conectarse vía SSH a un clúster de EMR. Para ello, establezca la configuración `EnableSSHLogin` en `true` en la configuración de seguridad de Amazon EMR que utilice para iniciar el clúster. Cuando se lance el clúster, use el siguiente comando para conectarse vía SSH en el clúster:

```
ssh username@EMR_PRIMARY_DNS_NAME
```

Tras ejecutar este comando, ingrese la contraseña de LDAP cuando se solicite.

Amazon EMR incluye un script en el clúster que permite a los usuarios generar un archivo keytab y un ticket de Kerberos para usarlos con aplicaciones compatibles que no aceptan credenciales del LDAP directamente. Algunas de estas aplicaciones incluyen `spark-submit` Spark SQL y. PySpark

Ejecute `ldap-kinit` y siga las instrucciones. Si la autenticación se realiza correctamente, el archivo keytab de Kerberos aparecerá en su directorio principal con un ticket de Kerberos válido. Utilice el ticket de Kerberos para ejecutar aplicaciones como lo haría en cualquier entorno kerberizado.

Integre Amazon EMR con AWS IAM Identity Center

Con las versiones 6.15.0 y posteriores de Amazon EMR, puede utilizar las identidades de AWS IAM Identity Center para autenticarse con un clúster de Amazon EMR. En las siguientes páginas, se proporciona información general conceptual, los requisitos previos y los pasos para lanzar un clúster de EMR con la integración de Identity Center.

Temas

- Descripción general
- Características y ventajas
- Introducción a la AWS IAM Identity Center integración para Amazon EMR
- Consideraciones y limitaciones de Amazon EMR con la integración de Identity Center

Descripción general

Identity Center es el enfoque recomendado para la autenticación y autorización de los empleados en organizaciones de cualquier AWS tamaño y tipo. Con Identity Center, puede crear y administrar identidades de usuario o conectar su fuente de identidad existente, que incluye Microsoft Active Directory, Okta, Ping Identity JumpCloud, Google Workspace y Microsoft Entra ID (anteriormente Azure AD). AWS

La propagación de identidades fiable es una AWS IAM Identity Center función que los administradores de Connected Servicios de AWS pueden utilizar para conceder y auditar el acceso a los datos del servicio. El acceso a estos datos se basa en los atributos del usuario, como las asociaciones de grupo. La configuración de una propagación de identidad fiable requiere la colaboración entre los administradores de Connected Servicios de AWS y los administradores del IAM Identity Center. Para obtener más información, consulte Requisitos previos y consideraciones.

Características y ventajas

La integración de Amazon EMR con IAM Identity Center ofrece los siguientes beneficios:

- Amazon EMR proporciona credenciales para transmitir una identidad de Identity Center a un clúster de EMR.
- Amazon EMR configura todas las aplicaciones compatibles para que se autentiquen con las credenciales del clúster.

- Amazon EMR configura y mantiene la seguridad de las aplicaciones compatibles con el protocolo de Kerberos, por lo que no se necesitan comandos ni scripts.
- La capacidad de aplicar la autorización a nivel de prefijo de Amazon S3 con las identidades de Identity Center en los prefijos de S3 administrados por S3 Access Grants.
- La capacidad de hacer cumplir la autorización a nivel de tabla con las identidades de Identity Center en las tablas AWS Lake Formation AWS Glue gestionadas.

Introducción a la AWS IAM Identity Center integración para Amazon EMR

Esta sección le ayuda a configurar Amazon EMR para que se integre con. AWS IAM Identity Center

Temas

- Creación de una instancia de Identity Center
- Creación de un rol de IAM para Identity Center
- Agregue permisos para los servicios que no están integrados con IAM Identity Center
- Creación de una configuración de seguridad habilitada por Identity Center
- Creación y lanzamiento de un clúster habilitado por Identity Center
- Configuración de Lake Formation para un clúster de EMR habilitado por IAM Identity Center
- Cómo trabajar con S3 Access Grants en un clúster de EMR habilitado por IAM Identity Center

> **Note**
>
> Para utilizar la integración de Identity Center con EMR, Lake Formation o Concesiones de acceso a S3 deben estar habilitados. También puede usar ambos. Si ninguna de las dos está habilitada, no se admite la integración de Identity Center.

Creación de una instancia de Identity Center

Si aún no tiene una instancia de Identity Center, cree una en la Región de AWS en la que desea lanzar el clúster de EMR. Una instancia de Identity Center solo puede existir en una sola región para una Cuenta de AWS.

Use el siguiente AWS CLI comando para crear una nueva instancia con el nombre*`MyInstance`*:

```
aws sso-admin create-instance --name MyInstance
```

Creación de un rol de IAM para Identity Center

Para integrar Amazon EMR con AWS IAM Identity Center, cree un rol de IAM que se autentique con Identity Center desde el clúster de EMR. Amazon EMR también utiliza credenciales SigV4 para transmitir la identidad de Identity Center a servicios posteriores, como AWS Lake Formation. El rol también debe tener los permisos correspondientes para invocar los servicios posteriores.

Al crear el rol, utilice la siguiente política de permisos:

```
{
  "Statement": [
    {
      "Sid": "IdCPermissions",
      "Effect": "Allow",
      "Action": [
        "sso-oauth:*"
      ],
      "Resource": "*"
    },
    {
      "Sid": "GlueandLakePermissions",
      "Effect": "Allow",
      "Action": [
        "glue:*",
        "lakeformation:GetDataAccess"
      ],
      "Resource": "*"
    },
    {
      "Sid": "AccessGrantsPermissions",
      "Effect": "Allow",
      "Action": [
        "s3:GetDataAccess",
        "s3:GetAccessGrantsInstanceForPrefix"
      ],
      "Resource": "*"
    }
  ]
}
```

La política de confianza de este rol permite que el rol InstanceProfile deje asumir el rol.

```
{
    "Sid": "AssumeRole",
    "Effect": "Allow",
    "Principal": {
        "AWS": "arn:aws:iam::12345678912:role/EMR_EC2_DefaultRole"
    },
    "Action": [
        "sts:AssumeRole",
        "sts:SetContext"
    ]
}
```

Si el rol no tiene credenciales de confianza y accede a una tabla protegida por Lake Formation, Amazon EMR establece automáticamente el `principalId` del rol asumido en *`userID`*-`untrusted`. El siguiente es un fragmento de un evento que muestra el. CloudTrail `principalId`

```
{
    "eventVersion": "1.09",
    "userIdentity": {
        "type": "AssumedRole",
        "principalId": "ABCDEFGH1JKLMNO2PQR3TU:5000-untrusted",
        "arn": "arn:aws:sts::123456789012:assumed-role/EMR_TIP/5000-untrusted",
        "accountId": "123456789012",
        "accessKeyId": "ABCDEFGH1IJKLMNOPQ7R3"
        ...
```

Agregue permisos para los servicios que no están integrados con IAM Identity Center

AWS credenciales que utilizan Trusted Identity Propagation: las políticas de IAM definidas en la función de IAM para cualquier llamada realizada a servicios que no estén integrados con el IAM Identity Center. Esto incluye, por ejemplo, el. AWS Key Management Service Su rol también debe definir los permisos de IAM para cualquiera de estos servicios a los que intente acceder. Los servicios integrados del Centro de Identidad de IAM compatibles actualmente incluyen AWS Lake Formation Amazon S3 Access Grants.

Para obtener más información sobre la propagación de identidad confiable, consulte Propagación de identidad confiable en todas las aplicaciones.

Creación de una configuración de seguridad habilitada por Identity Center

Para lanzar un clúster de EMR con la integración de IAM Identity Center, utilice el siguiente comando de ejemplo para crear una configuración de seguridad de Amazon EMR que tenga Identity Center habilitado. A continuación, se explica cada configuración.

```
aws emr create-security-configuration --name "IdentityCenterConfiguration-with-lf-
accessgrants" --region "us-west-2" --security-configuration '{
    "AuthenticationConfiguration":{
        "IdentityCenterConfiguration":{
            "EnableIdentityCenter":true,
            "IdentityCenterApplicationAssigmentRequired":false,
            "IdentityCenterInstanceARN": "arn:aws:sso:::instance/
ssoins-123xxxxxxxxx789"
        }
    },
    "AuthorizationConfiguration": {
        "LakeFormationConfiguration": {
            "AuthorizedSessionTagValue": "Amazon EMR"
        },
        "IAMConfiguration": {
          "EnableApplicationScopedIAMRole": true,
          "ApplicationScopedIAMRoleConfiguration": {
            "PropagateSourceIdentity": true
          }
        }
    },
    "EncryptionConfiguration": {
        "EnableInTransitEncryption": true,
        "EnableAtRestEncryption": false,
        "InTransitEncryptionConfiguration": {
            "TLSCertificateConfiguration": {
                "CertificateProviderType": "PEM",
                "S3Object": "s3://amzn-s3-demo-bucket/cert/my-certs.zip"
            }
        }
    }
}'
```

- **`EnableIdentityCenter`**: (obligatorio) habilita la integración de Identity Center.

- **`IdentityCenterInstanceARN`**: (obligatorio). El ARN de la instancia del Identity Center. Si no se incluye, se busca el ARN de la instancia del IAM Identity Center existente como parte del paso de configuración.
- **`IAMRoleForEMRIdentityCenterApplicationARN`**: (obligatorio) el rol de IAM que obtiene los tokens de Identity Center del clúster.
- **`IdentityCenterApplicationAssignmentRequired`** : (booleano) determina si será necesaria una asignación para utilizar la aplicación de Identity Center. Este campo es opcional. Si no se proporciona un valor, se utiliza el valor predeterminado `false`.
- **`AuthorizationConfiguration`** / **`LakeFormationConfiguration`**: si lo desea, configure la autenticación:
  - **`IAMConfiguration`**— Permite utilizar la función EMR Runtimes Roles además de su identidad TIP. Si habilita esta configuración, usted (o el AWS servicio que llama) deberán especificar un rol de tiempo de ejecución de IAM en cada llamada a EMR Steps o EMR. `GetClusterSessionCredentials` APIs Si el clúster de EMR se usa con SageMaker Unified Studio, esta opción es necesaria si la propagación de identidad confiable también está habilitada.
  - **`EnableLakeFormation`**: habilita la autorización de Lake Formation en el clúster.

Para habilitar la integración de Identity Center con Amazon EMR, debe especificar `EncryptionConfiguration` y `IntransitEncryptionConfiguration`.

Creación y lanzamiento de un clúster habilitado por Identity Center

Ahora que configuró el rol de IAM que se autentica con Identity Center y creó una configuración de seguridad de Amazon EMR que tiene Identity Center habilitado, puede crear y lanzar un clúster con reconocimiento de identidad. Para ver los pasos para lanzar el clúster con la configuración de seguridad requerida, consulte Especificar una configuración de seguridad para un clúster de Amazon EMR.

En las siguientes secciones, se describe cómo configurar el clúster habilitado por Identity Center con las opciones de seguridad compatibles con Amazon EMR:

- Cómo trabajar con S3 Access Grants en un clúster de EMR habilitado por IAM Identity Center
- Configuración de Lake Formation para un clúster de EMR habilitado por IAM Identity Center

Configuración de Lake Formation para un clúster de EMR habilitado por IAM Identity Center

Puede integrarlo AWS Lake Formationcon su clúster de EMR AWS IAM Identity Center habilitado.

En primer lugar, asegúrese de tener una instancia de Identity Center configurada en la misma región que el clúster. Para obtener más información, consulte Creación de una instancia de Identity Center. Puede encontrar el ARN de la instancia en la consola de IAM Identity Center en los detalles de la instancia o utilizar el siguiente comando para ver los detalles de todas las instancias desde la CLI:

```
aws sso-admin list-instances
```

A continuación, utilice el ARN y el ID de su AWS cuenta con el siguiente comando para configurar Lake Formation para que sea compatible con IAM Identity Center:

```
aws lakeformation create-lake-formation-identity-center-configuration --cli-input-json
 file://create-lake-fromation-idc-config.json
json input:
{
    "CatalogId": "account-id/org-account-id",
    "InstanceArn": "identity-center-instance-arn"
}
```

Ahora, llame a `put-data-lake-settings` y habilite `AllowFullTableExternalDataAccess` con Lake Formation:

```
aws lakeformation put-data-lake-settings --cli-input-json file://put-data-lake-
settings.json
json input:
{
    "DataLakeSettings": {
        "DataLakeAdmins": [
            {
                "DataLakePrincipalIdentifier": "admin-ARN"
            }
        ],
        "CreateDatabaseDefaultPermissions": [...],
        "CreateTableDefaultPermissions": [...],
        "AllowExternalDataFiltering": true,
        "AllowFullTableExternalDataAccess": true
```

```
    }
}
```

Por último, conceda permisos de tabla completos al ARN de identidad para el usuario que accede al clúster de EMR. El ARN contiene el ID de usuario de Identity Center. Vaya a Identity Center en la consola, seleccione Usuarios y, a continuación, seleccione el usuario para ver su configuración de Información general.

Copie el ID de usuario y péguelo en el siguiente ARN para *user-id*:

```
arn:aws:identitystore:::user/user-id
```

> **Note**
>
> Las consultas en el clúster de EMR solo funcionan si la identidad de IAM Identity Center tiene acceso total a la tabla protegida de Lake Formation. Si la identidad no tiene acceso total a la tabla, la consulta fallará.

Utilice el siguiente comando para conceder al usuario acceso total a la tabla:

```
aws lakeformation grant-permissions --cli-input-json file://grantpermissions.json
json input:
{
    "Principal": {
        "DataLakePrincipalIdentifier": "arn:aws:identitystore:::user/user-id"
    },
    "Resource": {
        "Table": {
            "DatabaseName": "tip_db",
            "Name": "tip_table"
        }
    },
    "Permissions": [
        "ALL"
    ],
    "PermissionsWithGrantOption": [
        "ALL"
    ]
}
```

Añadir la aplicación ARN a IDC para la integración de Lake Formation

Para consultar los recursos habilitados para Lake Formation, es necesario agregar el ARN de la aplicación de IDC. Para ello, sigue estos pasos:

1. En la consola, selecciona AWS Lake Formation.
2. Seleccione la integración de IAM Identity Center y la integración de aplicaciones Lake Formation haciendo coincidir el ARN de la aplicación. El ARN aparecerá en la lista de ID de la aplicación.

Cómo trabajar con S3 Access Grants en un clúster de EMR habilitado por IAM Identity Center

Puede integrar S3 Access Grants con su clúster de EMR AWS IAM Identity Center habilitado.

Utilice S3 Access Grants para autorizar el acceso a los conjuntos de datos desde los clústeres que utilizan Identity Center. Cree concesiones para aumentar los permisos que configura para usuarios, grupos o roles de IAM o para un directorio corporativo. Para obtener más información, consulte Uso de S3 Access Grants con Amazon EMR.

Temas

- Cómo crear una instancia y una ubicación de S3 Access Grants
- Creación de concesiones para las identidades de Identity Center

Cómo crear una instancia y una ubicación de S3 Access Grants

Si aún no tiene una instancia de S3 Access Grants, cree una en la Región de AWS en la que desea lanzar el clúster de EMR.

Use el siguiente AWS CLI comando para crear una nueva instancia con el nombre*MyInstance*:

```
aws s3control-access-grants create-access-grants-instance \
--account-id 12345678912 \
--identity-center-arn "identity-center-instance-arn" \
```

A continuación, cree una ubicación de S3 Access Grants y reemplace los valores en color rojo por los suyos:

```
aws s3control-access-grants create-access-grants-location \
```

```
--account-id 12345678912 \
--location-scope s3:// \
--iam-role-arn "access-grant-role-arn" \
--region aa-example-1
```

> **Note**
>
> Defina el parámetro `iam-role-arn` como ARN de `accessGrantRole`.

Creación de concesiones para las identidades de Identity Center

Por último, cree las concesiones para las identidades que tienen acceso al clúster:

```
aws s3control-access-grants create-access-grant \
--account-id 12345678912 \
--access-grants-location-id "default" \
--access-grants-location-configuration S3SubPrefix="s3-bucket-prefix"
--permission READ \
--grantee GranteeType=DIRECTORY_USER,GranteeIdentifier="your-identity-center-user-id"
```

Ejemplo de salida:

```
{
"CreatedAt": "2023-09-21T23:47:24.870000+00:00",
"AccessGrantId": "1234-12345-1234-1234567",
"AccessGrantArn": "arn:aws:s3:aa-example-1-1:123456789012:access-grants/default/grant/
xxxx1234-1234-5678-1234-1234567890",
"Grantee": {
"GranteeType": "DIRECTORY_USER",
"GranteeIdentifier": "5678-56789-5678-567890"
},
"AccessGrantsLocationId": "default",
"AccessGrantsLocationConfiguration": {
"S3SubPrefix": "myprefix/*"
},
"Permission": "READ",
"GrantScope": "s3://myprefix/*"
}
```

Consideraciones y limitaciones de Amazon EMR con la integración de Identity Center

Tenga en cuenta los siguientes puntos cuando utilice IAM Identity Center con Amazon EMR:

- La propagación de identidad confiable a través de Identity Center es compatible con Amazon EMR 6.15.0 y versiones posteriores, y solo con Apache Spark. Además, Amazon EMR 7.8.0 y versiones posteriores admiten la propagación confiable de identidades a través de Identity Center mediante la función EMR Runtime Roles, y solo con Apache Spark.
- Para habilitar los clústeres de EMR con propagación de identidades de confianza, debe usar AWS CLI para crear una configuración de seguridad que tenga habilitada la propagación de identidades de confianza y usar esa configuración de seguridad al lanzar el clúster. Para obtener más información, consulte Creación de una configuración de seguridad habilitada por Identity Center.
- Los controles de acceso detallados AWS Lake Formation que utilizan la propagación de identidad confiable están disponibles para los clústeres de Amazon EMR en la versión 7.2.0 y versiones posteriores de EMR. Entre las versiones 6.15.0 y 7.1.0 de EMR, solo está disponible el control de acceso a nivel de mesa, basado en Lake Formation. AWS
- Con los clústeres de Amazon EMR que utilizan Trusted Identity Propagation, las operaciones que admiten el control de acceso basado en Lake Formation con Apache Spark incluyen SELECT, ALTER TABLE, INSERT INTO y DROP TABLE.
- La propagación de identidad confiable con Amazon EMR se admite en los siguientes casos: Regiones de AWS
  - `af-south-1`: África (Ciudad del Cabo)
  - `ap-east-1`: Asia-Pacífico (Hong Kong)
  - `ap-northeast-1`: Asia Pacífico (Tokio)
  - `ap-northeast-2`: Asia-Pacífico (Seúl)
  - `ap-northeast-3`: Asia-Pacífico (Osaka)
  - `ap-south-1`: Asia-Pacífico (Bombay)
  - `ap-south-2`: Asia-Pacífico (Hyderabad)
  - `ap-southeast-1`: Asia Pacífico (Singapur)
  - `ap-southeast-2`: Asia Pacífico (Sídney)
  - `ap-southeast-3`: Asia-Pacífico (Yakarta)
  - `ap-southeast-4`: Asia-Pacífico (Melbourne)

- `ca-central-1`: Canadá (centro)
- `eu-central-1`: Europa (Fráncfort)
- `eu-central-2`: Europa (Zúrich)
- `eu-north-1`: Europa (Estocolmo)
- `eu-south-1`: Europa (Milán)
- `eu-south-2`: Europa (España)
- `eu-west-1`: Europa (Irlanda)
- `eu-west-2`: Europa (Londres)
- `eu-west-3`: Europa (París)
- `il-central-1`: Israel (Tel Aviv)
- `me-central-1`: Oriente-Medio (EAU)
- `me-south-1`: Medio Oriente (Baréin)
- `sa-east-1`: América del Sur (São Paulo)
- `us-east-1`: Este de EE. UU. (Norte de Virginia)
- `us-east-2`: Este de EE. UU. (Ohio)
- `us-west-1`: Oeste de EE. UU. (Norte de California)
- `us-west-2`: Oeste de EE. UU. (Oregón)

Integre Amazon EMR con AWS Lake Formation

AWS Lake Formation es un servicio gestionado que le ayuda a descubrir, catalogar, limpiar y proteger los datos de un lago de datos de Amazon Simple Storage Service (S3). Lake Formation proporciona un acceso detallado a nivel de columnas a las bases de datos y tablas del catálogo de datos de Glue. AWS Para obtener más información, consulte ¿Qué es AWS Lake Formation?

Con las versiones 6.7.0 y posteriores de Amazon EMR, puede aplicar un control de acceso basado en Lake Formation a los trabajos de Spark, Hive y Presto que envíe a los clústeres de Amazon EMR. Para integrar con Lake Formation, debe crear un clúster de EMR con un rol en tiempo de ejecución. Un rol en tiempo de ejecución es un rol de AWS Identity and Access Management (IAM) que se asocia a los trabajos o consultas de Amazon EMR. A continuación, Amazon EMR utiliza esta función para acceder AWS a los recursos. Para obtener más información, consulte Roles en tiempo de ejecución para los pasos de Amazon EMR.

Cómo funciona Amazon EMR con Lake Formation

Tras integrar Amazon EMR con Lake Formation, puede ejecutar consultas a los clústeres de Amazon EMR con la `StepAPI` o con AI Studio. SageMaker Luego, Lake Formation proporciona acceso a los datos a través de credenciales temporales para Amazon EMR. Este proceso se denomina “expedición de credenciales”. Para obtener más información, consulte ¿Qué es AWS Lake Formation?

A continuación, se ofrece una descripción general de alto nivel sobre cómo Amazon EMR obtiene acceso a los datos protegidos por las políticas de seguridad de Lake Formation.

1. Un usuario envía una solicitud de datos de Amazon EMR en Lake Formation.
2. Amazon EMR solicita credenciales temporales a Lake Formation para permitir que el usuario acceda a los datos.
3. Lake Formation devuelve credenciales temporales.
4. Amazon EMR envía la solicitud de consulta para obtener datos de Amazon S3.
5. Amazon EMR recibe los datos de Amazon S3, los filtra y devuelve los resultados en función de los permisos de usuario que el usuario definió en Lake Formation.

Para obtener más información sobre cómo agregar usuarios y grupos a las políticas de Lake Formation, consulte Concesión de permisos para el catálogo de datos.

Requisitos previos

Si desea integrar Amazon EMR y Lake Formation, debe cumplir los siguientes requisitos:

- Active la autorización de roles en tiempo de ejecución en el clúster de Amazon EMR.
- Utilice el catálogo de datos de AWS Glue como almacén de metadatos.
- Defina y gestione los permisos en Lake Formation para acceder a las bases de datos, tablas y columnas de AWS Glue Data Catalog. Para obtener más información, consulte ¿Qué es AWS Lake Formation?

Temas

- Habilitación de Lake Formation con Amazon EMR

- Uso de Apache Hudi y Lake Formation con Amazon EMR
- Apache Iceberg y Lake Formation con Amazon EMR
- Delta Lake y Lake Formation en Amazon EMR
- Consideraciones sobre Amazon EMR con Lake Formation
- Uso de vistas del catálogo de datos de Glue en Amazon EMR
- Uso de un rol de tiempo de ejecución de tareas de Lake Formation con acceso completo a las tablas

Habilitación de Lake Formation con Amazon EMR

Con Amazon EMR 6.15.0 y versiones posteriores, cuando ejecutas trabajos de Spark en Amazon EMR en EC2 clústeres que acceden a los datos del catálogo de datos de AWS Glue, puedes utilizarlos AWS Lake Formation para aplicar permisos a nivel de tabla, fila, columna y celda en tablas basadas en Hudi, Iceberg o Delta Lake.

En esta sección, explicamos cómo crear una configuración de seguridad y cómo configurar Lake Formation para que funcione con Amazon EMR. También explicamos cómo lanzar un clúster con la configuración de seguridad que creó para Lake Formation.

Paso 1: configuración de un rol en tiempo de ejecución para el clúster de EMR

Para usar un rol en tiempo de ejecución para el clúster de EMR, debe crear una configuración de seguridad. Con una configuración de seguridad, puede aplicar opciones consistentes de seguridad, autorización y autenticación en todos sus clústeres.

1. Cree un archivo denominado `lf-runtime-roles-sec-cfg.json` con la siguiente configuración de seguridad.

```
{
    "AuthorizationConfiguration": {
        "IAMConfiguration": {
            "EnableApplicationScopedIAMRole": true,
            "ApplicationScopedIAMRoleConfiguration": {
                "PropagateSourceIdentity": true
            }
        },
        "LakeFormationConfiguration": {
            "AuthorizedSessionTagValue": "Amazon EMR"
```

```
        }
    },
    "EncryptionConfiguration": {
     "EnableAtRestEncryption": false,
            "EnableInTransitEncryption": true,
            "InTransitEncryptionConfiguration": {
            "TLSCertificateConfiguration": {<certificate-configuration>}
        }
    }
}
```

El siguiente ejemplo ilustra cómo utilizar un archivo zip con certificados en Amazon S3 para la configuración de certificados:

- Como proveedor de claves se utiliza un archivo zip con certificados en Amazon S3. (Consulte Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR los requisitos de los certificados).

```
"TLSCertificateConfiguration": {
 "CertificateProviderType": "PEM",
 "S3Object": "s3://MyConfigStore/artifacts/MyCerts.zip"
 }
```

El siguiente ejemplo ilustra cómo utilizar un proveedor de claves personalizado para la configuración del certificado:

- Se utiliza un proveedor de claves personalizado. (Consulte Proporcionar certificados para el cifrado de datos en tránsito con el cifrado de Amazon EMR los requisitos de certificación).

```
"TLSCertificateConfiguration": {
 "CertificateProviderType": "Custom",
 "S3Object": "s3://MyConfig/artifacts/MyCerts.jar",
 "CertificateProviderClass": "com.mycompany.MyCertProvider"
                        }
```

2. A continuación, para asegurarse de que la etiqueta de sesión puede autorizar Lake Formation, establezca la propiedad `LakeFormationConfiguration/AuthorizedSessionTagValue` en `Amazon EMR`.
3. Use el siguiente comando para crear la configuración de seguridad de Amazon EMR.

```
aws emr create-security-configuration \
--name 'iamconfig-with-iam-lf' \
--security-configuration file://lf-runtime-roles-sec-cfg.json
```

Como alternativa, puede utilizar la consola de Amazon EMR para crear una configuración de seguridad con ajustes personalizados.

Paso 2: lanzar un clúster de Amazon EMR

Ahora tiene todo listo para lanzar un clúster de EMR con la configuración de seguridad que creó en el paso anterior. Para obtener más información sobre las configuraciones de seguridad, consulte Uso de configuraciones de seguridad para configurar la seguridad del clúster de Amazon EMR y Roles en tiempo de ejecución para los pasos de Amazon EMR.

Paso 3a: configurar permisos a nivel de tabla basados en Lake Formation con los roles de tiempo de ejecución de Amazon EMR

Si no necesita un control de acceso detallado a nivel de columna, fila o celda, puede configurar permisos a nivel de tabla con el catálogo de datos de Glue. Para habilitar el acceso a nivel de tabla, vaya a la AWS Lake Formation consola y seleccione la opción de configuración de integración de aplicaciones en la sección Administración de la barra lateral. Luego, habilite la siguiente opción y elija Guardar:

Permitir que motores externos accedan a datos en las ubicaciones de Amazon S3 con acceso total a las tablas

Paso 3b: configurar permisos a nivel de columna, fila o celda basados en Lake Formation con los roles de tiempo de ejecución de Amazon EMR

Para aplicar permisos a nivel de tabla y columna con Lake Formation, el administrador del lago de datos de Lake Formation debe establecer `Amazon EMR` como valor para la configuración de la etiqueta de sesión, `AuthorizedSessionTagValue`. Lake Formation usa esta etiqueta de sesión para autorizar a los intermediarios y proporcionar acceso al lago de datos. Puede configurar esta etiqueta de sesión en la sección Filtrado de datos externos de la consola de Lake Formation. Sustitúyala por tu *`123456789012`* propia ID. Cuenta de AWS

Paso 4: Configurar las subvenciones de AWS Glue and Lake Formation para las funciones de tiempo de ejecución de Amazon EMR

Para continuar con la configuración del control de acceso basado en Lake Formation con las funciones de tiempo de ejecución de Amazon EMR, debe configurar las concesiones de AWS Glue and Lake Formation para las funciones de ejecución de Amazon EMR. Para permitir que sus roles en tiempo de ejecución de IAM interactúen con Lake Formation, debe concederles acceso con `lakeformation:GetDataAccess` y`glue:Get*`.

Los permisos de Lake Formation controlan el acceso a los recursos del catálogo de datos de AWS Glue, a las ubicaciones de Amazon S3 y a los datos subyacentes en esas ubicaciones. Los permisos de la IAM controlan el acceso a Lake Formation and AWS Glue APIs y a los recursos. Aunque es posible que tenga el permiso de Lake Formation para acceder a una tabla del catálogo de datos (SELECT), la operación fallará si no tiene el permiso de IAM en la API `glue:Get*`. Para obtener más información sobre el control de acceso de Lake Formation, consulte Información general sobre el control de acceso de Lake Formation.

1. Cree el archivo `emr-runtime-roles-lake-formation-policy.json` con el siguiente contenido.

   JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "LakeFormationManagedAccess",
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess",
        "glue:Get*",
        "glue:Create*",
        "glue:Update*"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
}
```

2. Cree la política de IAM relacionada.

```
aws iam create-policy \
--policy-name emr-runtime-roles-lake-formation-policy \
--policy-document file://emr-runtime-roles-lake-formation-policy.json
```

3. Para asignar esta política a sus roles en tiempo de ejecución de IAM, siga los pasos que se indican en Administración de permisos de AWS Lake Formation.

Ahora puede usar roles en tiempo de ejecución y Lake Formation para aplicar permisos de tabla y columna. También puede utilizar una identidad de origen para controlar las acciones y supervisar las AWS CloudTrail operaciones.

Para cada rol de IAM que vaya a utilizar como rol en tiempo de ejecución, defina la siguiente política de confianza y sustituya `EMR_EC2_DefaultRole` por el rol de perfil de instancia. Para modificar la política de confianza de un rol de IAM, consulte Modificación de una política de confianza de rol.

```
{
   "Sid":"AllowAssumeRole",
   "Effect":"Allow",
   "Principal":{
     "AWS":"arn:aws:iam::<AWS_ACCOUNT_ID>:role/EMR_EC2_DefaultRole"
   },
   "Action":[
        "sts:AssumeRole",
        "sts:TagSession"
       ]
 }
```

Para ver un end-to-end ejemplo detallado, consulte Introducción a las funciones de tiempo de ejecución para los pasos de Amazon EMR.

Para obtener información sobre cómo integrarse con Iceberg y AWS Glue Data Catalog para una jerarquía de varios catálogos, consulte Configurar Spark para acceder a una jerarquía de varios catálogos en AWS Glue Data Catalog.

Uso de Apache Hudi y Lake Formation con Amazon EMR

Las versiones 6.15.0 y posteriores de Amazon EMR incluyen compatibilidad con un control de acceso detallado basado en Apache Hudi al leer y AWS Lake Formation escribir datos con Spark

SQL. Amazon EMR es compatible con el control de acceso a nivel de tabla, fila, columna y celda con Apache Hudi. Con esta función, puede ejecutar consultas de instantáneas en copy-on-write las tablas para consultar la última instantánea de la tabla en un instante de confirmación o compactación determinado.

Actualmente, un clúster de Amazon EMR habilitado para Lake Formation debe recuperar la columna de tiempo de confirmación de Hudi para realizar consultas incrementales y consultas de viajes en el tiempo. No es compatible con la sintaxis `timestamp as of` de Spark ni con la función `Spark.read()`. La sintaxis correcta es `select * from table where _hoodie_commit_time <= point_in_time`. Para obtener más información, consulte la tabla Consultas sobre viajes en el tiempo puntuales en Hudi.

La siguiente matriz de compatibilidad enumera algunas de las características principales de Apache Hudi con Lake Formation:

	Copiar al escribir	fusionar al leer
Consultas de instantáneas: Spark SQL	✓	✓
Consultas optimizadas para la lectura: Spark SQL	✓	✓
Consultas incrementales	✓	✓
Consultas de viaje en el tiempo	✓	✓
Tabla de metadatos	✓	✓
Comandos **`INSERT`** DML	✓	✓
Comandos DDL		
Consultas de orígenes de datos de Spark		
Escrituras de orígenes de datos de Spark		

Consulta de tablas de Hudi

En esta sección, se muestra cómo ejecutar las consultas compatibles descritas anteriormente en un clúster habilitado para Lake Formation. La tabla debe ser una tabla de catálogo registrada.

1. Para iniciar el intérprete de comandos de Spark, utilice los siguientes comandos.

```
spark-sql
--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,com.amazonaws.em
 \
--conf spark.sql.catalog.spark_catalog.lf.managed=true
```

   Si quiere que Lake Formation utilice el servidor de registros para administrar su catálogo de Spark, establezca `spark.sql.catalog.<managed_catalog_name>.lf.managed` en verdadero.

2. Para consultar la última instantánea de copy-on-write las tablas, usa los siguientes comandos.

```
SELECT * FROM my_hudi_cow_table
```

```
spark.read.table("my_hudi_cow_table")
```

3. Para conocer los datos compactados más recientes de las tablas `MOR`, puede consultar la tabla optimizada para lectura que tiene el sufijo `_ro`:

```
SELECT * FROM my_hudi_mor_table_ro
```

```
spark.read.table("my_hudi_mor_table_ro")
```

> **Note**
>
> El rendimiento de las lecturas en los clústeres de Lake Formation puede ser más lento debido a las optimizaciones que no se admiten. Estas características incluyen la lista de

archivos basada en los metadatos de Hudi y la omisión de datos. Recomendamos probar el rendimiento de su aplicación para asegurarse de que cumple con los requisitos.

Apache Iceberg y Lake Formation con Amazon EMR

Las versiones 6.15.0 y posteriores de Amazon EMR incluyen compatibilidad con un control de acceso detallado basado en Apache Iceberg al leer y AWS Lake Formation escribir datos con Spark SQL. Amazon EMR es compatible con el control de acceso a nivel de tabla, fila, columna y celda con Apache Iceberg. Con esta función, puede ejecutar consultas de instantáneas en copy-on-write las tablas para consultar la última instantánea de la tabla en un instante de confirmación o compactación determinado.

Si desea utilizar el formato Iceberg, defina las siguientes configuraciones. Reemplace *DB_LOCATION* por la ruta de Amazon S3 en la que se encuentran las tablas de Iceberg y reemplace los marcadores de posición de Región e ID de cuenta por sus propios valores.

```
spark-sql \
--conf
 spark.sql.extensions=org.apache.iceberg.spark.extensions.IcebergSparkSessionExtensions,com.ama

--conf spark.sql.catalog.iceberg_catalog=org.apache.iceberg.spark.SparkCatalog
--conf spark.sql.catalog.iceberg_catalog.warehouse=s3://DB_LOCATION
--conf spark.sql.catalog.iceberg_catalog.catalog-
impl=org.apache.iceberg.aws.glue.GlueCatalog
--conf spark.sql.catalog.iceberg_catalog.io-impl=org.apache.iceberg.aws.s3.S3FileIO
--conf spark.sql.catalog.iceberg_catalog.glue.account-id=ACCOUNT_ID
--conf spark.sql.catalog.iceberg_catalog.glue.id=ACCOUNT_ID
--conf spark.sql.catalog.iceberg_catalog.client.assume-role.region=AWS_REGION
--conf spark.sql.secureCatalog=iceberg_catalog
```

Si quiere que Lake Formation utilice el servidor de registros para administrar su catálogo de Spark, establezca `spark.sql.catalog.<managed_catalog_name>.lf.managed` en verdadero.

También debe tener cuidado de NO pasar por alto la siguiente configuración de roles:

```
--conf spark.sql.catalog.my_catalog.client.assume-role.region
--conf spark.sql.catalog.my_catalog.client.assume-role.arn
```

```
--conf spark.sql.catalog.my_catalog.client.assume-
role.tags.LakeFormationAuthorizedCaller
```

La siguiente matriz de compatibilidad enumera algunas de las características principales de Apache Iceberg con Lake Formation:

	Copiar al escribir	fusionar al leer
Consultas de instantáneas: Spark SQL	✓	✓
Consultas optimizadas para la lectura: Spark SQL	✓	✓
Consultas incrementales	✓	✓
Consultas de viaje en el tiempo	✓	✓
Tabla de metadatos	✓	✓
Comandos **`INSERT`** DML	✓	✓
Comandos DDL		
Consultas de orígenes de datos de Spark		
Escrituras de orígenes de datos de Spark		

Delta Lake y Lake Formation en Amazon EMR

Las versiones 6.15.0 y posteriores de Amazon EMR incluyen compatibilidad con un control de acceso detallado basado en Delta Lake al leer y AWS Lake Formation escribir datos con Spark SQL. Amazon EMR es compatible con el control de acceso a nivel de tabla, fila, columna y celda con Delta Lake. Con esta función, puede ejecutar consultas de instantáneas en copy-on-write las tablas para consultar la última instantánea de la tabla en un instante de confirmación o compactación determinado.

Para usar Delta Lake con Lake Formation, ejecute el siguiente comando.

```
spark-sql \
--conf
 spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension,com.amazonaws.emr.recordserver.co
 \
--conf spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog
 \
--conf spark.sql.catalog.spark_catalog.lf.managed=true
```

Si quiere que Lake Formation utilice el servidor de registros para administrar su catálogo de Spark, establezca `spark.sql.catalog.<managed_catalog_name>.lf.managed` en verdadero.

La siguiente matriz de compatibilidad enumera algunas de las características principales de Delta Lake con Lake Formation:

	Copiar al escribir	fusionar al leer
Consultas de instantáneas: Spark SQL	✓	✓
Consultas optimizadas para la lectura: Spark SQL	✓	✓
Consultas incrementales	No admitido	No admitido
Consultas de viaje en el tiempo	No admitido	No admitido
Tabla de metadatos	✓	✓
Comandos **`INSERT`** DML	✓	✓
Comandos DDL		
Consultas de orígenes de datos de Spark		
Escrituras de orígenes de datos de Spark		

Creación de una tabla de Delta Lake en AWS Glue Data Catalog

Amazon EMR with Lake Formation no admite los comandos DDL ni la creación de tablas Delta. Siga estos pasos para crear tablas en el catálogo de datos de AWS Glue.

1. Cree una tabla Delta con el siguiente ejemplo. Asegúrese de que su ubicación de S3 exista.

```
spark-sql \
--conf "spark.sql.extensions=io.delta.sql.DeltaSparkSessionExtension" \
--conf
 "spark.sql.catalog.spark_catalog=org.apache.spark.sql.delta.catalog.DeltaCatalog"

> CREATE DATABASE if not exists <DATABASE_NAME> LOCATION 's3://<S3_LOCATION>/
transactionaldata/native-delta/<DATABASE_NAME>/';
> CREATE TABLE <TABLE_NAME> (x INT, y STRING, z STRING) USING delta;
> INSERT INTO <TABLE_NAME> VALUES (1, 'a1', 'b1');
```

2. Para ver los detalles de la tabla, vaya a https://console.aws.amazon.com/glue/.
3. En el menú de navegación de la izquierda, expanda el Catálogo de datos, elija Tablas y, a continuación, elija la tabla que ha creado. En Esquema, deberías ver que la tabla Delta que creaste con Spark almacena todas las columnas en un tipo de datos llamado `array<string>` in AWS Glue.
4. Para definir filtros a nivel de columna y celda en Lake Formation, elimine la columna `col` del esquema y, a continuación, agregue las columnas que están en el esquema de la tabla. En este ejemplo, añada las columnas `x`, `y` y `z`.

Consideraciones sobre Amazon EMR con Lake Formation

Tenga en cuenta lo siguiente cuando utilice Amazon EMR con. AWS Lake Formation

Amazon EMR with Lake Formation está disponible en todas las regiones disponibles.

- Amazon EMR admite un control de acceso detallado a través de Lake Formation solo para las tablas Apache Hive y Apache Iceberg. Los formatos de Apache Hive incluyen Parquet, ORC y xSV.
- No puede desactivar `DynamicResourceAllocation` para los trabajos de Lake Formation.
- Solo puede utilizar Lake Formation con trabajos de Spark.
- Amazon EMR con Lake Formation solo admite una sola sesión de Spark durante un trabajo.

- Amazon EMR con Lake Formation solo admite consultas de tablas entre cuentas compartidas a través de enlaces de recursos.
- Lo siguiente no es compatible:
  - Conjuntos de datos distribuidos resilientes (RDD)
  - Streaming de Spark
  - Lectura con permisos concedidos de Lake Formation
  - Control de acceso para columnas anidadas
- Amazon EMR bloquea las funcionalidades que podrían socavar el aislamiento total del controlador del sistema, incluidas las siguientes:
  - UDTs, Hive UDFs y cualquier función definida por el usuario que incluya clases personalizadas
  - Orígenes de datos personalizados
  - Suministro de archivos jar adicionales para la extensión, el conector o el metaalmacén de Spark
  - `ANALYZE TABLE` command
- Para hacer cumplir los controles de acceso, `EXPLAIN PLAN` y las operaciones de DDL, como `DESCRIBE TABLE`, no exponen información restringida.
- Amazon EMR restringe el acceso a los registros de Spark del controlador del sistema en las aplicaciones habilitadas para Lake Formation. Dado que el controlador del sistema se ejecuta con permisos elevados, los eventos y registros que genera el controlador del sistema pueden incluir información confidencial. Para evitar que usuarios o códigos no autorizados accedan a estos datos confidenciales, Amazon EMR deshabilita el acceso a los registros de controladores del sistema.

  Los registros de los perfiles del sistema siempre se conservan en el almacenamiento gestionado; se trata de una configuración obligatoria que no se puede desactivar. Estos registros se almacenan de forma segura y se cifran mediante una clave de KMS gestionada por el cliente o una clave de KMS AWS gestionada.

  Si su aplicación de Amazon EMR se encuentra en una subred privada con puntos de enlace de VPC para Amazon S3 y adjunta una política de puntos de enlace para controlar el acceso, antes de que sus trabajos puedan enviar datos de registro a AWS Amazon S3 gestionado, debe incluir los permisos detallados en Almacenamiento gestionado en su política de VPC al punto de enlace de puerta de enlace de S3. Para solicitudes de solución de problemas, póngase en contacto con el soporte. AWS
- Si ha registrado una ubicación de tabla en Lake Formation, la ruta de acceso a los datos pasa por las credenciales almacenadas de Lake Formation, independientemente del permiso de IAM para la función de ejecución de tareas de Amazon EMR. Si configura incorrectamente el rol registrado con

la ubicación de la tabla, los trabajos enviados que usen el rol con permisos de IAM de S3 para la ubicación de la tabla fallarán.

- Para escribir en una tabla de Lake Formation se utiliza el permiso de IAM en lugar de los permisos concedidos por Lake Formation. Si el rol de tiempo de ejecución de su trabajo tiene los permisos de S3 necesarios, puede usarlo para ejecutar operaciones de escritura.

A continuación, se indican las consideraciones y limitaciones cuando se utiliza Apache Iceberg:

- Solo puede usar Apache Iceberg con el catálogo de sesiones y no con catálogos con nombres arbitrarios.
- Las tablas de Iceberg que están registradas en Lake Formation solo admiten las tablas de metadatos `history`, `metadata_log_entries`, `snapshots`, `files`, `manifests` y `refs`. Amazon EMR oculta las columnas que pueden contener datos confidenciales, como `partitions`, `path` y `summaries`. Esta limitación no se aplica a las tablas de Iceberg que no estén registradas en Lake Formation.
- Las tablas que no se registran en Lake Formation admiten todos los procedimientos almacenados de Iceberg. Los procedimientos `register_table` y `migrate` no son compatibles con ninguna tabla.
- Le recomendamos que utilice Iceberg DataFrameWriter V2 en lugar de V1.
- EMR 7.10 proporciona una forma de volver a RecordServer utilizar funciones compatibles con el FGAC nativo RecordServer, pero que aún no lo son, como la reescritura en las tablas registradas de Lake Formation. Para volver a la versión anterior, especifique las siguientes configuraciones al lanzar el clúster.

```
{
    "Classification": "spark-defaults",
    "Properties": {
      "spark.emr.lakeformation.legacy.enabled": "true"
    }
  },
  {
    "Classification": "yarn-site",
    "Properties": {
      "spark.emr.lakeformation.legacy.enabled": "true"
    }
  }
```

Uso de vistas del catálogo de datos de Glue en Amazon EMR

> **Note**
>
> La creación y administración de vistas del catálogo de datos de AWS Glue para su uso con EMR EC2 está disponible en Amazon EMR versión 7.10.0 y versiones posteriores.

Puede crear y gestionar vistas en el catálogo de datos de AWS Glue para utilizarlas con la EMR activada. EC2 Estas se conocen comúnmente como vistas del catálogo de datos de AWS Glue. Estas vistas son útiles porque admiten varios motores de consultas SQL, por lo que puede acceder a la misma vista en distintos AWS servicios, como EMR on y EC2 Amazon Amazon Athena Redshift.

Al crear una vista en el catálogo de datos, puede utilizar las concesiones de recursos y los controles de acceso basados en etiquetas AWS Lake Formation para conceder el acceso a ella. Con este método de control de acceso, no tiene que configurar el acceso adicional a las tablas a las que hizo referencia en el momento de crear la vista. Este método de concesión de permisos se denomina semántica del definidor y estas vistas se denominan vistas del definidor. Para obtener más información sobre el control de acceso en Lake Formation, consulte Concesión y revocación de permisos sobre los recursos del catálogo de datos en la Guía para desarrolladores de AWS Lake Formation.

Las vistas del Catálogo de datos son útiles para los siguientes casos de uso:

- Control de acceso detallado: puede crear una vista que restrinja el acceso a los datos en función de los permisos que necesite el usuario. Por ejemplo, puede usar las vistas del catálogo de datos para evitar que el personal que no trabaje en el departamento de Recursos Humanos (RR.HH.) vea información de identificación personal (PII).
- Definición de vista completa: al aplicar filtros a la vista del catálogo de datos, se asegura de que los registros de datos disponibles en una vista del catálogo de datos estén siempre completos.
- Seguridad mejorada: la definición de consulta utilizada para crear la vista debe estar completa. Esta ventaja significa que las vistas del catálogo de datos son menos susceptibles a los comandos SQL de actores malintencionados.
- Compartir datos de forma sencilla: comparta datos con otras AWS cuentas sin moverlos. Para obtener más información, consulte Uso compartido entre cuentas en Lake Formation.

Creación de una vista del catálogo de datos

Existen diferentes formas de crear una vista de catálogo de datos. Estas incluyen el uso de Spark SQL AWS CLI o Spark. A continuación se muestran algunos ejemplos.

Using SQL

A continuación se muestra la sintaxis para crear una vista de catálogo de datos. Anote el tipo de `MULTI DIALECT` vista. Esto distingue la vista del catálogo de datos de otras vistas. El `SECURITY` predicado se especifica como. `DEFINER` Esto indica una vista del catálogo de datos con `DEFINER` semántica.

```
CREATE [ OR REPLACE ] PROTECTED MULTI DIALECT VIEW [IF NOT EXISTS] view_name
[(column_name [COMMENT column_comment], ...) ]
[ COMMENT view_comment ]
[TBLPROPERTIES (property_name = property_value, ... )]
SECURITY DEFINER
AS query;
```

El siguiente es un ejemplo de `CREATE` declaración, siguiendo la sintaxis:

```
CREATE PROTECTED MULTI DIALECT VIEW catalog_view
SECURITY DEFINER
AS
SELECT order_date, sum(totalprice) AS price
FROM source_table
GROUP BY order_date
```

También puede crear una vista en modo de ejecución en seco, mediante SQL, para probar la creación de la vista, sin crear realmente el recurso. El uso de esta opción da como resultado una «ejecución en seco» que valida la entrada y, si la validación se realiza correctamente, devuelve el JSON del objeto de tabla AWS Glue que representará la vista. En este caso, no se crea la vista real.

```
CREATE [ OR REPLACE ] PROTECTED MULTI DIALECT VIEW view_name
SECURITY DEFINER
[ SHOW VIEW JSON ]
AS view-sql
```

Using the AWS CLI

> **Note**
>
> Cuando se utiliza el comando CLI, el SQL utilizado para crear la vista no se analiza. Esto puede provocar que se cree la vista, pero las consultas no se realicen correctamente. Asegúrese de probar la sintaxis de SQL antes de crear la vista.

Para crear una vista, utilice el siguiente comando CLI:

```
aws glue create-table --cli-input-json '{
  "DatabaseName": "database",
  "TableInput": {
    "Name": "view",
    "StorageDescriptor": {
      "Columns": [
        {
          "Name": "col1",
          "Type": "data-type"
        },
        ...
        {
          "Name": "col_n",
          "Type": "data-type"
        }
      ],
      "SerdeInfo": {}
    },
    "ViewDefinition": {
      "SubObjects": [
        "arn:aws:glue:aws-region:aws-account-id:table/database/referenced-table1",
        ...
        "arn:aws:glue:aws-region:aws-account-id:table/database/referenced-tableN",
       ],
      "IsProtected": true,
      "Representations": [
        {
          "Dialect": "SPARK",
          "DialectVersion": "1.0",
          "ViewOriginalText": "Spark-SQL",
          "ViewExpandedText": "Spark-SQL"
```

```
        }
      ]
    }
  }
}'
```

Operaciones de vista admitidas

Los siguientes fragmentos de comandos muestran varias formas de trabajar con las vistas del Catálogo de datos:

- CREATE VIEW

  Crea una vista del Catálogo de datos. El siguiente es un ejemplo que muestra cómo crear una vista a partir de una tabla existente:

  ```
  CREATE PROTECTED MULTI DIALECT VIEW catalog_view
  SECURITY DEFINER AS SELECT * FROM my_catalog.my_database.source_table
  ```

- ALTER VIEW

  Sintaxis disponible:

  - `ALTER VIEW view_name [FORCE] ADD DIALECT AS query`
  - `ALTER VIEW view_name [FORCE] UPDATE DIALECT AS query`
  - `ALTER VIEW view_name DROP DIALECT`

  Puede utilizar la opción `FORCE ADD DIALECT` para forzar la actualización del esquema y los subobjetos según el nuevo dialecto del motor. Tenga en cuenta que hacer esto puede provocar errores de consulta si no utiliza también `FORCE` para actualizar otros dialectos del motor. A continuación se muestra un ejemplo:

  ```
  ALTER VIEW catalog_view FORCE ADD DIALECT
  AS
  SELECT order_date, sum(totalprice) AS price
  FROM source_table
  GROUP BY orderdate;
  ```

  A continuación se muestra cómo modificar una vista para actualizar el dialecto:

```
ALTER VIEW catalog_view UPDATE DIALECT AS
SELECT count(*) FROM my_catalog.my_database.source_table;
```

- DESCRIBE VIEW

  Sintaxis disponible para describir una vista:

  - `SHOW COLUMNS {FROM|IN} view_name [{FROM|IN} database_name]`— Si el usuario tiene los permisos de AWS Glue and Lake Formation necesarios para describir la vista, puede enumerar las columnas. A continuación se muestran un par de comandos de ejemplo para mostrar columnas:

    ```
    SHOW COLUMNS FROM my_database.source_table;
    SHOW COLUMNS IN my_database.source_table;
    ```

  - `DESCRIBE view_name`— Si el usuario tiene los permisos de AWS Glue and Lake Formation necesarios para describir la vista, puede enumerar las columnas de la vista junto con sus metadatos.

- DROP VIEW

  Sintaxis disponible:

  - `DROP VIEW [ IF EXISTS ] view_name`

    El siguiente ejemplo muestra una instrucción `DROP` que comprueba si existe una vista antes de eliminarla:

    ```
    DROP VIEW IF EXISTS catalog_view;
    ```

- MOSTRAR, CREAR VISTA

  - `SHOW CREATE VIEW view_name`: Muestra la instrucción SQL que crea la vista especificada. El siguiente es un ejemplo que muestra cómo crear una vista en el catálogo de datos:

    ```
    SHOW CREATE TABLE my_database.catalog_view;
    CREATE PROTECTED MULTI DIALECT VIEW my_catalog.my_database.catalog_view (
      net_profit,
      customer_id,
      item_id,
      sold_date)
    TBLPROPERTIES (
      'transient_lastDdlTime' = '1736267222')
    ```

```
SECURITY DEFINER AS SELECT * FROM
my_database.store_sales_partitioned_lf WHERE customer_id IN (SELECT customer_id
 from source_table limit 10)
```

- SHOW VIEWS

  Muestra todas las vistas del catálogo, como las vistas normales, las vistas multidialectales (MDV) y las MDV sin el dialecto de Spark. La sintaxis disponible es la siguiente:

  - `SHOW VIEWS [{ FROM | IN } database_name] [LIKE regex_pattern]`:

    A continuación se muestra un comando de ejemplo para mostrar vistas:

    ```
    SHOW VIEWS IN marketing_analytics LIKE 'catalog_view*';
    ```

Para obtener más información sobre la creación y configuración de vistas de catálogos de datos, consulte Building AWS Glue Data Catalog views en la Guía para AWS Lake Formation desarrolladores.

Consulta de la vista del Catálogo de datos

Tras crear una vista del catálogo de datos, puede consultarla mediante un trabajo de Amazon EMR Spark que tenga activado el control de AWS Lake Formation acceso detallado. El rol de ejecución del trabajo debe tener el `SELECT` permiso Lake Formation en la vista del catálogo de datos. No es necesario conceder acceso a las tablas subyacentes a las que se hace referencia en la vista.

Una vez que haya configurado todo, podrá consultar la vista. Por ejemplo, después de crear una aplicación Amazon EMR en EMR Studio, puede ejecutar la siguiente consulta para acceder a una vista.

```
SELECT * from my_database.catalog_view LIMIT 10;
```

Una función útil es la. `invoker_principal` Devuelve el identificador único de la función de ejecución del trabajo de EMRS. Esto se puede usar para controlar el resultado de la vista, en función del principio que lo invoca. Puede utilizarla para añadir una condición a la vista que defina los resultados de la consulta en función del rol que realiza la llamada. El rol de ejecución del trabajo debe tener permiso para la acción de `LakeFormation:GetDataLakePrincipal` IAM para utilizar esta función.

```
select invoker_principal();
```

Puede añadir esta función a una `WHERE` cláusula, por ejemplo, para refinar los resultados de la consulta.

Consideraciones y limitaciones

Al crear vistas del catálogo de datos, se aplica lo siguiente:

- Solo puede crear vistas del catálogo de datos con Amazon EMR 7.10 y versiones posteriores.
- El creador de la vista en el Catálogo de datos debe tener acceso a `SELECT` a las tablas base subyacentes a las que accede la vista. La creación de la vista en el Catálogo de datos falla si una tabla base específica tiene filtros de Lake Formation impuestos al rol del creador.
- Las tablas base no deben tener el permiso `IAMAllowedPrincipals` de lago de datos en Lake Formation. Si está presente, se produce el error «Las vistas en varios dialectos solo pueden hacer referencia a tablas sin permisos de IAMAllowed director».
- La ubicación de Amazon S3 de la tabla debe estar registrada como ubicación de lago de datos de Lake Formation. Si la tabla no está registrada, se produce el error de que las vistas en varios dialectos solo pueden hacer referencia a las tablas gestionadas por Lake Formation. Para obtener información sobre cómo registrar ubicaciones de Amazon S3 en Lake Formation, consulte Registrar una ubicación de Amazon S3 en la Guía para AWS Lake Formation desarrolladores.
- Solo puede crear vistas del catálogo de datos `PROTECTED`. No se admiten las vistas `UNPROTECTED`.
- No puede hacer referencia a las tablas de otra AWS cuenta en una definición de vista del catálogo de datos. Tampoco se puede hacer referencia a una tabla de la misma cuenta que esté en una región distinta.
- Para compartir datos en una cuenta o región, toda la vista debe compartirse entre cuentas y regiones, utilizando los enlaces de recursos de Lake Formation.
- No se admiten las funciones definidas por el usuario (UDFs).
- Puede utilizar vistas basadas en tablas Iceberg. También se admiten los formatos de tabla abierta Apache Hudi y Delta Lake.
- No puede hacer referencia a otras vistas en las vistas del catálogo de datos.
- Un esquema de vista del catálogo de datos de AWS Glue siempre se almacena en minúsculas. Por ejemplo, si utiliza una sentencia DDL para crear una vista del catálogo de datos de Glue con una columna denominada`Castle`, la columna creada en el catálogo de datos de Glue aparecerá

en minúsculas, a. `castle` Si, a continuación, especifica el nombre de la columna en una consulta de DML como `Castle` o`CASTLE`, EMR Spark pondrá el nombre en minúsculas para poder ejecutar la consulta. Sin embargo, el encabezado de la columna se muestra con mayúsculas y minúsculas que especificó en la consulta.

Si desea que una consulta falle en el caso de que el nombre de una columna especificado en la consulta de DML no coincida con el nombre de la columna del catálogo de datos de Glue, puede `spark.sql.caseSensitive=true` configurarlo.

Uso de un rol de tiempo de ejecución de tareas de Lake Formation con acceso completo a las tablas

Los temas de esta sección tratan sobre cómo puede acceder a las tablas del catálogo de Glue Data protegidas por Lake Formation desde trabajos de EMR Spark o sesiones interactivas con acceso completo a las tablas.

Acceso sin filtros a Lake Formation para EMR en EC2

Con las versiones 7.8.0 y posteriores de Amazon EMR, puede aprovechar AWS Lake Formation con Glue Data Catalog, donde el rol de ejecución de tareas tiene permisos de tabla completos. Esta capacidad le permite leer y escribir en tablas protegidas por Lake Formation desde sus trabajos interactivos y por lotes de EMR Spark.

Uso de Lake Formation con acceso completo a la mesa

Puede acceder a las tablas del catálogo de Glue Data protegidas por AWS Lake Formation desde trabajos de EMR Spark o sesiones interactivas en las que el rol de tiempo de ejecución del trabajo tiene acceso completo a las tablas. No es necesario que habilite AWS Lake Formation en el clúster de EMR. Cuando un trabajo de Spark está configurado para el acceso total a la tabla (FTA), las credenciales de AWS Lake Formation se utilizan para leer y escribir datos de S3 para las tablas registradas de AWS Lake Formation, mientras que las credenciales del rol de tiempo de ejecución del trabajo se utilizan para read/write las tablas no registradas en Lake Formation AWS .

> **Note**
>
> Para ejecutar Full Table Access (FTA), no se requiere ni se admite un control de acceso detallado de Lake Formation para un trabajo.

> **Note**
>
> El acceso completo a la tabla no requiere el uso de una función de tiempo de ejecución de EMR y funciona tanto con la función de tiempo de ejecución como con la función de servicio para las instancias de clúster EC2 . Para obtener más información, consulte los pasos Rol de servicio para EC2 instancias de clúster (perfil de EC2 instancia) o Roles de tiempo de ejecución para Amazon EMR.

Paso 1: Habilitar el acceso completo a la mesa en Lake Formation

Para usar el modo Full Table Access (FTA), debe permitir que los motores de consulta de terceros accedan a los datos sin la validación de etiquetas de sesión de IAM en AWS Lake Formation. Para habilitarlo, siga los pasos descritos en Integración de aplicaciones para acceso completo a la tabla .

> **Note**
>
> Al acceder a las tablas entre cuentas, se debe habilitar el acceso completo a las tablas tanto en la cuenta del productor como en la del consumidor. Del mismo modo, al acceder a las tablas entre regiones, esta configuración debe estar habilitada tanto en las regiones de productores como de consumidores.

Paso 2: configuración de los permisos de IAM para el rol en tiempo de ejecución del trabajo

Para el acceso de lectura o escritura a los datos subyacentes, además de los permisos de Lake Formation, un rol de ejecución de tareas necesita el permiso de `lakeformation:GetDataAccess` IAM. Con este permiso, Lake Formation concede la solicitud de credenciales temporales para acceder a los datos.

El siguiente es un ejemplo de política sobre cómo proporcionar permisos de IAM para acceder a un script en Amazon S3, cargar registros en S3, permisos de la API AWS Glue y permiso para acceder a Lake Formation.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
```

```
    {
      "Sid": "ScriptAccess",
      "Effect": "Allow",
      "Action": [
        "s3:GetObject"
      ],
      "Resource": [
        "arn:aws:s3:::*.amzn-s3-demo-bucket/scripts"
      ]
    },
    {
      "Sid": "LoggingAccess",
      "Effect": "Allow",
      "Action": [
        "s3:PutObject"
      ],
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket/logs/*"
      ]
    },
    {
      "Sid": "GlueCatalogAccess",
      "Effect": "Allow",
      "Action": [
        "glue:Get*",
        "glue:Create*",
        "glue:Update*"
      ],
      "Resource": [
        "*"
      ]
    },
    {
      "Sid": "LakeFormationAccess",
      "Effect": "Allow",
      "Action": [
        "lakeformation:GetDataAccess"
      ],
      "Resource": [
        "*"
      ]
    }
  ]
```

```
}
```

Paso 2.1 Configurar los permisos de Lake Formation

- Los trabajos de Spark que leen datos de S3 requieren el permiso Lake Formation SELECT.
- Los trabajos de Spark write/delete cuyos datos en S3 requieren el permiso de Lake Formation ALL (SUPER).
- Los trabajos de Spark que interactúan con el catálogo de Glue Data requieren los permisos DESCRIBIR, ALTER y DROP, según corresponda.

Para obtener más información, consulte Otorgar permisos sobre los recursos del catálogo de datos.

Paso 3: Inicializa una sesión de Spark para tener acceso completo a la mesa mediante Lake Formation

Requisitos previos

AWS El catálogo de datos de Glue debe configurarse como un metaalmacén para acceder a las tablas de Lake Formation.

Defina los siguientes ajustes para configurar el catálogo de Glue como un metaalmacén:

```
--conf spark.sql.catalogImplementation=hive
--conf
 spark.hive.metastore.client.factory.class=com.amazonaws.glue.catalog.metastore.AWSGlueDataCata
```

Para activar el catálogo de datos para EMR EC2, consulta Cómo usar AWS Glue Data Catalog con Spark en Amazon EMR.

Acceso a las tablas

Para acceder a las tablas registradas en AWS Lake Formation, se deben establecer las siguientes configuraciones durante la inicialización de Spark para configurar Spark para que use las credenciales de AWS Lake Formation.

Hive

```
--conf
 spark.hadoop.fs.s3.credentialsResolverClass=com.amazonaws.glue.accesscontrol.AWSLakeFormati
--conf spark.hadoop.fs.s3.useDirectoryHeaderAsFolderObject=true
```

```
--conf spark.hadoop.fs.s3.folderObject.autoAction.disabled=true
--conf spark.sql.catalog.skipLocationValidationOnCreateTable.enabled=true
--conf spark.sql.catalog.createDirectoryAfterTable.enabled=true
--conf spark.sql.catalog.dropDirectoryBeforeTable.enabled=true
```

Iceberg

```
--conf spark.sql.catalog.spark_catalog=org.apache.iceberg.spark.SparkSessionCatalog
--conf spark.sql.catalog.spark_catalog.warehouse=S3_DATA_LOCATION
--conf spark.sql.catalog.spark_catalog.client.region=REGION
--conf spark.sql.catalog.spark_catalog.type=glue
--conf spark.sql.catalog.spark_catalog.glue.account-id=ACCOUNT_ID
--conf spark.sql.catalog.spark_catalog.glue.lakeformation-enabled=true
--conf spark.sql.catalog.dropDirectoryBeforeTable.enabled=true
```

- `spark.hadoop.fs.s3.credentialsResolverClass=com.amazonaws.glue.accesscontrol.AW` Configure el sistema de archivos EMR (EMRFS) para usar las credenciales S3 de Lake Formation para las tablas registradas de AWS Lake Formation. Si la tabla no está registrada, utilice las credenciales del rol en tiempo de ejecución del trabajo.
- `spark.hadoop.fs.s3.useDirectoryHeaderAsFolderObject=trueyspark.hadoop.fs.s3.fol` Configure EMRFS para que utilice el encabezado del tipo de contenido application/x-directory en lugar del sufijo al crear carpetas S3. `$folder$` Esto es obligatorio al leer las tablas de Lake Formation, ya que las credenciales de Lake Formation no permiten leer carpetas de tablas con el sufijo $folder$.
- `spark.sql.catalog.skipLocationValidationOnCreateTable.enabled=true`: configure Spark para omitir la validación de que la ubicación de la tabla esté vacía antes de crearla. Esto es necesario para las tablas registradas de Lake Formation, ya que las credenciales de Lake Formation para verificar la ubicación vacía solo están disponibles después de crear la tabla de Glue Data Catalog. Sin esta configuración, las credenciales del rol en tiempo de ejecución del trabajo validarán si la ubicación de la tabla está vacía.
- `spark.sql.catalog.createDirectoryAfterTable.enabled=true`: configure Spark para crear la carpeta en Amazon S3 después de crear la tabla en el repositorio de metadatos de Hive. Esto es obligatorio para las tablas registradas de Lake Formation, ya que las credenciales de Lake Formation para crear la carpeta S3 solo están disponibles después de crear la tabla de Glue Data Catalog.
- `spark.sql.catalog.dropDirectoryBeforeTable.enabled=true`: Configura Spark para que elimine la carpeta S3 en el metaalmacén de Hive antes de eliminar la tabla. Esto es necesario

para las tablas registradas de Lake Formation, ya que las credenciales de Lake Formation para eliminar la carpeta S3 no están disponibles después de eliminar la tabla del catálogo de datos de Glue.

- `spark.sql.catalog.<catalog>.glue.lakeformation-enabled=true`: Configure el catálogo de Iceberg para usar las credenciales S3 de AWS Lake Formation para las tablas registradas de Lake Formation. Si la tabla no está registrada, utilice las credenciales predeterminadas del entorno.

Configure el modo de acceso completo a las tablas en SageMaker Unified Studio

Para acceder a las tablas registradas de Lake Formation desde las sesiones interactivas de Spark en JupyterLab cuadernos, debes usar el modo de permiso de compatibilidad. Usa el comando `%%configure` mágico para configurar tu configuración de Spark. Elija la configuración según el tipo de tabla:

For Hive tables

```
%%configure -f
{
    "conf": {
        "spark.hadoop.fs.s3.credentialsResolverClass":
 "com.amazonaws.glue.accesscontrol.AWSLakeFormationCredentialResolver",
        "spark.hadoop.fs.s3.useDirectoryHeaderAsFolderObject": true,
        "spark.hadoop.fs.s3.folderObject.autoAction.disabled": true,
        "spark.sql.catalog.skipLocationValidationOnCreateTable.enabled": true,
        "spark.sql.catalog.createDirectoryAfterTable.enabled": true,
        "spark.sql.catalog.dropDirectoryBeforeTable.enabled": true
    }
}
```

For Iceberg tables

```
%%configure -f
{
    "conf": {
        "spark.sql.catalog.spark_catalog":
 "org.apache.iceberg.spark.SparkSessionCatalog",
        "spark.sql.catalog.spark_catalog.warehouse": "S3_DATA_LOCATION",
        "spark.sql.catalog.spark_catalog.client.region": "REGION",
  "conf spark.sql.catalog.spark_catalog.type": "glue",
```

```
        "spark.sql.catalog.spark_catalog.glue.account-id": "ACCOUNT_ID",
        "spark.sql.catalog.spark_catalog.glue.lakeformation-enabled": true,
        "spark.sql.catalog.dropDirectoryBeforeTable.enabled": true
    }
}
```

Reemplace los marcadores de posición:

- `S3_DATA_LOCATION`: La ruta de tu bucket de S3
- `REGION`: AWS región (por ejemplo, us-east-1)
- `ACCOUNT_ID`: El ID de tu cuenta AWS

> **Note**
>
> Debe establecer estas configuraciones antes de ejecutar cualquier operación de Spark en el bloc de notas.

Operaciones admitidas

Estas operaciones utilizarán las credenciales de AWS Lake Formation para acceder a los datos de la tabla.

- CREATE TABLE
- ALTER TABLE
- INSERT INTO
- INSERT OVERWRITE
- UPDATE
- MERGE INTO
- DELETE FROM
- ANALIZAR TABLA
- REPARAR TABLA
- DROP TABLE
- Consultas de orígenes de datos de Spark

- Escrituras de orígenes de datos de Spark

> **Note**
>
> Las operaciones no mencionadas anteriormente aún usan permisos de IAM para acceder a los datos de las tablas.

Consideraciones

- Si se crea una tabla de Hive con un trabajo que no tiene habilitado el acceso total a la tabla y no se inserta ningún registro, las lecturas o escrituras posteriores de un trabajo con acceso total a la tabla fallarán. Esto se debe a que EMR Spark sin acceso completo a la tabla agrega el `$folder$` sufijo al nombre de la carpeta de la tabla. Para resolver esto, puede optar por una de las siguientes acciones:
  - Insertar al menos una fila en la tabla desde un trabajo que no tenga FTA habilitado.
  - Configure el trabajo que no tiene habilitado el FTA para que no utilice el `$folder$` sufijo en el nombre de la carpeta en S3. Esto se logra al configurar `spark.hadoop.fs.s3.useDirectoryHeaderAsFolderObject=true` en Spark.
  - Cree una carpeta S3 en la ubicación de la tabla `s3://path/to/table/table_name` mediante la consola AWS S3 o la CLI de AWS S3.
- El acceso completo a la tabla (FTA) funciona exclusivamente con el sistema de archivos de EMR (EMRFS). El sistema de archivos de S3A no es compatible.
- El acceso completo a tablas (FTA) es compatible con tablas de Hive e Iceberg. Aún no se ha agregado compatibilidad con tablas Hudi y Delta.
- Los trabajos que hagan referencia a tablas con las reglas de control de acceso detallado de Lake Formation (FGAC) o las vistas del catálogo de datos de Glue fallarán.
- El acceso completo a tablas no ofrece compatibilidad con Spark Streaming.
- Al DataFrame escribir Spark en una tabla de Lake Formation, solo se admite el modo APPEND: `df.write.mode("append").saveAsTable(table_name)`

Integración de Amazon EMR con Apache Ranger

A partir de Amazon EMR 5.32.0, puede iniciar un clúster que se integre de forma nativa con Apache Ranger. Apache Ranger es un marco de código abierto para habilitar, supervisar y administrar

la seguridad integral de los datos en toda la plataforma Hadoop. Para obtener más información, consulte Apache Ranger. Con la integración nativa, puede utilizar su propio Apache Ranger para aplicar un control de acceso a los datos detallado en Amazon EMR.

Esta sección proporciona información general sobre la integración de Amazon EMR con Apache Ranger. También incluye los requisitos previos y los pasos necesarios para lanzar un clúster de Amazon EMR integrado con Apache Ranger.

La integración nativa de Amazon EMR con Apache Ranger ofrece los siguientes beneficios clave:

- Control de acceso detallado a las bases de datos y tablas del metaalmacén de Hive, que le permite definir políticas de filtrado de datos en bases de datos, tablas y columnas para las aplicaciones Apache Spark y Apache Hive. Las aplicaciones de Hive admiten el filtrado de filas y el enmascaramiento de datos.
- La posibilidad de utilizar sus políticas de Hive existentes directamente con las aplicaciones de Amazon EMR para Hive.
- Control de acceso a los datos de prefijos y objetos de Amazon S3, lo que le permite definir políticas de filtrado de datos para acceder a los datos de S3 mediante el sistema de archivos de EMR.
- La capacidad de utilizar los CloudWatch registros para realizar auditorías centralizadas.
- Amazon EMR instala y administra los complementos de Apache Ranger en su nombre.

Apache Ranger con Amazon EMR

Apache Ranger es un marco para habilitar, supervisar y administrar la seguridad integral de los datos en toda la plataforma Hadoop.

Apache Ranger tiene las siguientes características:

- Administración de seguridad centralizada para gestionar todas las tareas relacionadas con la seguridad en una interfaz de usuario central o mediante REST APIs.
- Autorización detallada para realizar una acción u operación específica con un componente o herramienta de Hadoop, administrada a través de una herramienta de administración central.
- Un método de autorización estandarizado para todos los componentes de Hadoop.
- Compatibilidad mejorada para varios métodos de autorización.
- Auditoría centralizada del acceso de los usuarios y de las acciones administrativas (relacionadas con la seguridad) en todos los componentes de Hadoop.

Apache Ranger utiliza dos componentes clave para la autorización:

- Servidor de administración de políticas de Apache Ranger: este servidor le permite definir las políticas de autorización para las aplicaciones de Hadoop. Al llevar a cabo la integración con Amazon EMR, podrá definir y aplicar políticas para que Apache Spark y Hive accedan al metaalmacén de Hive, así como para acceder al Sistema de archivos de EMR (EMRFS) de datos de Amazon S3. Puede configurar un servidor de administración de políticas de Apache Ranger nuevo o utilizar uno existente para integrarlo con Amazon EMR.
- Complemento de Apache Ranger: este complemento valida el acceso de un usuario según las políticas de autorización definidas en el servidor de administración de políticas de Apache Ranger. Amazon EMR instala y configura el complemento Apache Ranger automáticamente para cada aplicación de Hadoop seleccionada en la configuración de Apache Ranger.

Temas

- Arquitectura de la integración de Amazon EMR con Apache Ranger
- Componentes de Amazon EMR para su uso con Apache Ranger

Arquitectura de la integración de Amazon EMR con Apache Ranger

Componentes de Amazon EMR para su uso con Apache Ranger

Amazon EMR permite obtener un control integral del acceso con Apache Ranger mediante los siguientes componentes. Consulte el diagrama de arquitectura para ver una representación visual de estos componentes de Amazon EMR con los complementos de Apache Ranger.

Agente secreto: el agente secreto almacena secretos de forma segura y los distribuye a otros componentes o aplicaciones de Amazon EMR. Los secretos pueden incluir credenciales de usuario temporales, claves de cifrado o tickets de Kerberos. El agente secreto se ejecuta en todos los nodos del clúster e intercepta las llamadas al servicio de metadatos de la instancia. Para las solicitudes de credenciales del rol del perfil de instancia, el agente secreto vende las credenciales en función del usuario solicitante y de los recursos solicitados tras autorizar la solicitud con el complemento S3 Ranger de EMRFS. El agente secreto se ejecuta como el *`emrsecretagent`*usuario y escribe los registros en el emr/secretagent/log directorio/. El proceso recurre a un conjunto específico de reglas `iptables` para funcionar. Es importante asegurarse de que `iptables` no esté deshabilitado. Si personaliza la configuración de `iptables`, las reglas de la tabla NAT deben conservarse y no modificarse.

Servidor de registros de EMR: el servidor de registros recibe las solicitudes de Spark para acceder a los datos. A continuación, autoriza las solicitudes reenviando los recursos solicitados al complemento Spark Ranger para Amazon EMR. El servidor de registros lee los datos de Amazon S3 y devuelve los datos filtrados a los que el usuario está autorizado a acceder según la política de Ranger. El servidor de registros se ejecuta en todos los nodos del clúster como usuario emr_record_server y escribe los registros en el directorio/-record-server. var/log/emr

Consideraciones sobre el uso de Amazon EMR con Apache Ranger

Aplicaciones compatibles con Amazon EMR con Apache Ranger

La integración entre Amazon EMR y Apache Ranger, en la que EMR instala los complementos de Ranger, actualmente admite las siguientes aplicaciones:

- Apache Spark (disponible con EMR 5.32+ y EMR 6.3+)
- Apache Hive (disponible con EMR 5.32+ y EMR 6.3+)
- Acceso a S3 a través de EMRFS (disponible con EMR 5.32+ y EMR 6.3+)

Las siguientes aplicaciones se pueden instalar en un clúster de EMR y es posible que deban configurarse para satisfacer sus necesidades de seguridad:

- Apache Hadoop (disponible con EMR 5.32+ y EMR 6.3+, incluidos YARN y HDFS)
- Apache Livy (disponible con EMR 5.32+ y EMR 6.3+)
- Apache Zeppelin (disponible con EMR 5.32+ y EMR 6.3+)
- Apache Hue (disponible con EMR 5.32+ y EMR 6.3+)
- Ganglia (disponible con EMR 5.32+ y EMR 6.3+)
- HCatalog (Disponible con EMR 5.32+ y EMR 6.3+)
- Mahout (disponible con EMR 5.32+ y EMR 6.3+)
- MXNet (Disponible con EMR 5.32+ y EMR 6.3+)
- TensorFlow (Disponible con EMR 5.32+ y EMR 6.3+)
- Tez (disponible con EMR 5.32+ y EMR 6.3+)
- Trino (disponible con EMR 6.7+)
- ZooKeeper (Disponible con EMR 5.32+ y EMR 6.3+)

> **Important**
>
> Las aplicaciones enumeradas anteriormente son las únicas aplicaciones compatibles actualmente. Para garantizar la seguridad del clúster, se le permite crear un clúster de EMR solo con las aplicaciones de la lista anterior cuando Apache Ranger esté habilitado. Actualmente no se admiten otras aplicaciones. Para garantizar la seguridad del clúster, si se intenta instalar otras aplicaciones, se rechazará el clúster.
> AWS No se admiten los formatos Glue Data Catalog y Open table, como Apache Hudi, Delta Lake y Apache Iceberg.

Funciones de Amazon EMR compatibles con Apache Ranger

Al utilizar Amazon EMR con Apache Ranger, se admiten las siguientes funciones de Amazon EMR:

- Cifrado en reposo y en tránsito
- Autenticación de Kerberos (obligatoria)
- Grupos de instancias, flotas de instancias e instancias de spot
- Reconfiguración de aplicaciones en un clúster en ejecución
- Cifrado del servidor (SSE) de EMRFS

> **Note**
>
> La configuración de cifrado de Amazon EMR rige el SSE. Para obtener más información, consulte Opciones de cifrado.

Limitaciones de la aplicación

Hay varias limitaciones que se deben tener en cuenta al integrar Amazon EMR y Apache Ranger:

- Actualmente, no puede utilizar la consola para crear una configuración de seguridad que especifique la opción de integración de AWS Ranger en el. AWS GovCloud (US) Region La configuración de seguridad se puede llevar a cabo con la CLI.
- Kerberos tiene que estar instalado en el clúster.

- Las aplicaciones UIs (interfaces de usuario), como la interfaz de usuario del administrador de recursos de YARN, la interfaz de usuario de HDFS y la NameNode interfaz de usuario de Livy, no están configuradas con la autenticación de forma predeterminada.
- Los permisos predeterminados de HDFS `umask` están configurados de forma que los objetos creados tengan el valor `world wide readable` de forma predeterminada.
- Amazon EMR no admite el modo de alta disponibilidad (varias entidades principales) con Apache Ranger.
- Para ver otras limitaciones, consulte las limitaciones de cada aplicación.

> **Note**
>
> La configuración de cifrado de Amazon EMR rige el SSE. Para obtener más información, consulte Opciones de cifrado.

Limitaciones del complemento

Cada complemento tiene limitaciones específicas. Para ver las limitaciones del complemento Apache Hive, consulte Apache Hive plugin limitations. Para ver las limitaciones del complemento Apache Spark, consulte Apache Spark plugin limitations. Para ver las limitaciones del complemento EMRFS S3, consulte EMRFS S3 plugin limitations.

Configuración de Amazon EMR para Apache Ranger

Antes de instalar Apache Ranger, revise la información de esta sección para asegurarse de que Amazon EMR esté correctamente configurado.

Temas

- Configure un servidor Ranger Admin para integrarlo con Amazon EMR
- Roles de IAM para la integración nativa con Apache Ranger
- Creación de la configuración de seguridad de EMR
- Almacenamiento de certificados TLS en AWS Secrets Manager
- Inicio de un clúster de EMR con Apache Ranger
- Configuración de Zeppelin para clústeres de Amazon EMR habilitados para Apache Ranger
- Problemas conocidos para la integración de Amazon EMR

Configure un servidor Ranger Admin para integrarlo con Amazon EMR

Para la integración de Amazon EMR, los complementos de la aplicación Apache Ranger deben comunicarse con el servidor de administración mediante TLS/SSL.

Requisito previo: habilitación del protocolo SSL del servidor de Ranger Admin

Apache Ranger en Amazon EMR requiere una comunicación SSL bidireccional entre los complementos y el servidor de Ranger Admin. Para garantizar que los complementos se comuniquen con el servidor Apache Ranger a través de SSL, habilite el siguiente atributo dentro de ranger-admin-site .xml en el servidor de administración de Ranger.

```
<property>
    <name>ranger.service.https.attrib.ssl.enabled</name>
    <value>true</value>
</property>
```

Además, se requieren las siguientes configuraciones:

```
<property>
    <name>ranger.https.attrib.keystore.file</name>
    <value>_<PATH_TO_KEYSTORE>_</value>
</property>

<property>
    <name>ranger.service.https.attrib.keystore.file</name>
    <value>_<PATH_TO_KEYSTORE>_</value>
</property>

<property>
    <name>ranger.service.https.attrib.keystore.pass</name>
    <value>_<KEYSTORE_PASSWORD>_</value>
</property>

<property>
    <name>ranger.service.https.attrib.keystore.keyalias</name>
    <value><PRIVATE_CERTIFICATE_KEY_ALIAS></value>
</property>

<property>
    <name>ranger.service.https.attrib.clientAuth</name>
    <value>want</value>
```

```
</property>

<property>
    <name>ranger.service.https.port</name>
    <value>6182</value>
</property>
```

Certificados TLS para la integración de Apache Ranger con Amazon EMR

La integración de Apache Ranger con Amazon EMR requiere que el tráfico desde los nodos de Amazon EMR al servidor de Ranger Admin esté cifrado mediante TLS y que los complementos de Ranger se autentiquen en el servidor de Apache Ranger mediante una autenticación TLS mutua bidireccional. El servicio Amazon EMR necesita el certificado público de su servidor de Ranger Admin (especificado en el ejemplo anterior) y el certificado privado.

Certificados del complemento Apache Ranger

El servidor de Apache Ranger Admin debe poder acceder a los certificados TLS públicos del complemento Apache Ranger para validarlos cuando se conecten los complementos. Hay tres métodos diferentes para hacerlo.

Método 1: configuración de un almacén de confianza en el servidor de Apache Ranger Admin

Rellene las siguientes configuraciones en ranger-admin-site .xml para configurar un almacén de confianza.

```
<property>
    <name>ranger.truststore.file</name>
    <value><LOCATION TO TRUSTSTORE></value>
</property>

<property>
    <name>ranger.truststore.password</name>
    <value><PASSWORD FOR TRUSTSTORE></value>
</property>
```

Método 2: carga del certificado en el almacén de confianza de certificados CA (cacerts) de Java

Si el servidor de Ranger Admin no especifica un almacén de confianza en sus opciones de JVM, puede colocar los certificados públicos del complemento en el almacén de cacerts predeterminado.

Método 3: creación de un almacén de confianza y configuración como parte de las opciones de JVM

En `{RANGER_HOME_DIRECTORY}/ews/ranger-admin-services.sh`, modifique `JAVA_OPTS` para que incluya `"-Djavax.net.ssl.trustStore=<TRUSTSTORE_LOCATION>"` y `"-Djavax.net.ssl.trustStorePassword=<TRUSTSTORE_PASSWORD>"`. Por ejemplo, agregue la siguiente línea después de la variable JAVA_OPTS existente.

```
JAVA_OPTS=" ${JAVA_OPTS} -Djavax.net.ssl.trustStore=${RANGER_HOME}/truststore/
truststore.jck -Djavax.net.ssl.trustStorePassword=changeit"
```

> **Note**
>
> Esta especificación puede exponer la contraseña del almacén de confianza si algún usuario puede iniciar sesión en el servidor de Apache Ranger Admin y ver los procesos en ejecución, por ejemplo, cuando utiliza el comando `ps`.

Uso de certificados autofirmados

Los certificados autofirmados no se recomiendan como certificados. Los certificados autofirmados no se pueden revocar y los certificados autofirmados pueden no cumplir con los requisitos de seguridad internos.

Instalación de definición de servicio para la integración de Ranger con Amazon EMR

El servidor de Ranger Admin utiliza una definición de servicio para describir los atributos de las políticas de una aplicación. Luego, las políticas se almacenan en un repositorio de políticas para que los clientes las descarguen.

Para poder configurar las definiciones de servicio, las llamadas REST se deben realizar al servidor de Ranger Admin. Consulte Apache Ranger Public APIsv2 para ver si es APIs obligatorio en la siguiente sección.

Instalación de la definición de servicio de Apache Spark

Para instalar la definición de servicio de Apache Spark, consulte Complemento de Apache Spark para la integración de Ranger con Amazon EMR.

Instalación de la definición de servicio de EMRFS

Para instalar la definición de servicio de S3 para Amazon EMR, consulte Complemento EMRFS S3 para la integración de Ranger con Amazon EMR.

Uso de la definición de servicio de Hive

Apache Hive puede usar la definición de servicio de Ranger existente que se incluye con Apache Ranger 2.0 y versiones posteriores. Para obtener más información, consulte Complemento de Apache Hive para la integración de Ranger con Amazon EMR.

Reglas de tráfico de red para la integración con Amazon EMR

Cuando Apache Ranger se integra con su clúster de EMR, el clúster debe comunicarse con servidores adicionales y AWS.

Todos los nodos de Amazon EMR, incluidos los nodos principales y de tareas, deben poder comunicarse con los servidores de Apache Ranger Admin para descargar las políticas. Si su administrador de Apache Ranger se ejecuta en Amazon EC2, debe actualizar el grupo de seguridad para poder captar el tráfico del clúster de EMR.

Además de comunicarse con el servidor Ranger Admin, todos los nodos deben poder comunicarse con los siguientes servicios: AWS

- Amazon S3
- AWS KMS (si usa EMRFS SSE-KMS)
- Amazon CloudWatch
- AWS STS

Si planea ejecutar su clúster de EMR en una subred privada, configure la VPC para que pueda comunicarse con estos servicios mediante AWS PrivateLink y puntos de conexión de VPC en la Guía del usuario de Amazon VPC o mediante una instancia de traducción de direcciones de red (NAT) de la Guía del usuario de Amazon VPC.

Roles de IAM para la integración nativa con Apache Ranger

La integración entre Amazon EMR y Apache Ranger se basa en tres roles clave que debe crear antes de lanzar el clúster:

- Un perfil de EC2 instancia de Amazon personalizado para Amazon EMR
- Un rol de IAM para los motores de Apache Ranger
- Una función de IAM para otros servicios AWS

Esta sección proporciona información general sobre estos roles y las políticas que debe incluir en cada rol de IAM. Para obtener más información sobre la creación de estos roles, consulte Configure un servidor Ranger Admin para integrarlo con Amazon EMR.

EC2 perfil de instancia para Amazon EMR

Amazon EMR utiliza un rol de servicio de IAM para realizar acciones en su nombre para aprovisionar y administrar clústeres. La función de servicio para EC2 las instancias de clúster, también denominada perfil de EC2 instancia para Amazon EMR, es un tipo especial de función de servicio que se asigna a cada EC2 instancia de un clúster en el momento del lanzamiento.

Para definir los permisos para la interacción del clúster de EMR con los datos de Amazon S3 y con el metaalmacén de Hive protegido por Apache Ranger y otros AWS servicios, defina un perfil de EC2 instancia personalizado para usarlo en lugar del que `EMR_EC2_DefaultRole` se utiliza al lanzar el clúster.

Para obtener más información, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia) y Personalización de roles de IAM con Amazon EMR.

Debe añadir las siguientes instrucciones al perfil de EC2 instancia predeterminado de Amazon EMR para poder etiquetar las sesiones y acceder al AWS Secrets Manager que almacena los certificados TLS.

```
    {
      "Sid": "AllowAssumeOfRolesAndTagging",
      "Effect": "Allow",
      "Action": ["sts:TagSession", "sts:AssumeRole"],
      "Resource": [
        "arn:aws:iam::<AWS_ACCOUNT_ID>:role/<RANGER_ENGINE-
PLUGIN_DATA_ACCESS_ROLE_NAME>",
        "arn:aws:iam::<AWS_ACCOUNT_ID>:role/<RANGER_USER_ACCESS_ROLE_NAME>"
      ]
    },
    {
        "Sid": "AllowSecretsRetrieval",
        "Effect": "Allow",
        "Action": "secretsmanager:GetSecretValue",
        "Resource": [

 "arn:aws:secretsmanager:<REGION>:<AWS_ACCOUNT_ID>:secret:<PLUGIN_TLS_SECRET_NAME>*",

 "arn:aws:secretsmanager:<REGION>:<AWS_ACCOUNT_ID>:secret:<ADMIN_RANGER_SERVER_TLS_SECRET_NAME>
```

```
        ]
    }
```

> **Note**
>
> Para los permisos de Secrets Manager, no olvide el comodín (“*”) que aparece al final del nombre del secreto o sus solicitudes fallarán. El comodín es para las versiones de los secretos.

> **Note**
>
> Limite el alcance de la AWS Secrets Manager política a solo los certificados necesarios para el aprovisionamiento.

Rol de IAM para Apache Ranger

Este rol proporciona credenciales para que los motores de ejecución de confianza, como Apache Hive y Servidor de registros de Amazon EMR, puedan acceder a los datos de Amazon S3. Utilice únicamente este rol para acceder a los datos de Amazon S3, incluidas las claves de KMS, si utiliza S3 SSE-KMS.

Este rol debe crearse con la política mínima que se indica en el siguiente ejemplo.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Sid": "CloudwatchLogsPermissions",
      "Action": [
        "logs:CreateLogGroup",
        "logs:CreateLogStream",
        "logs:PutLogEvents"
      ],
      "Effect": "Allow",
      "Resource": [
```

```
        "arn:aws:logs:*:123456789012:log-
group:CLOUDWATCH_LOG_GROUP_NAME_IN_SECURITY_CONFIGURATION:*"
      ]
    },
    {
      "Sid": "BucketPermissionsInS3Buckets",
      "Action": [
        "s3:CreateBucket",
        "s3:DeleteBucket",
        "s3:ListAllMyBuckets",
        "s3:ListBucket"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket1",
        "arn:aws:s3:::amzn-s3-demo-bucket2"
      ]
    },
    {
      "Sid": "ObjectPermissionsInS3Objects",
      "Action": [
        "s3:GetObject",
        "s3:DeleteObject",
        "s3:PutObject"
      ],
      "Effect": "Allow",
      "Resource": [
        "arn:aws:s3:::amzn-s3-demo-bucket1/*",
        "arn:aws:s3:::amzn-s3-demo-bucket2/*"
      ]
    }
  ]
}
```

Important

Se debe incluir el asterisco «*» al final del recurso de CloudWatch registro para permitir escribir en los flujos de registro.

> **Note**
>
> Si utiliza la visión de consistencia de EMRFS o el cifrado S3-SSE, agregue permisos a las tablas de DynamoDB y a las claves de KMS para que los motores de ejecución puedan interactuar con esos motores.

El rol de IAM de Apache Ranger lo asume el rol de perfil de EC2 instancia. Utilice el siguiente ejemplo para crear una política de confianza que permita que el rol de perfil de instancia asuma el rol de IAM de Apache Ranger. EC2

```
    {
      "Sid": "",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::<AWS_ACCOUNT_ID>:role/<EC2 INSTANCE PROFILE ROLE NAME eg.
EMR_EC2_DefaultRole>"
      },
      "Action": ["sts:AssumeRole", "sts:TagSession"]
    }
```

Rol de IAM para otros servicios AWS para integración de Amazon EMR

Esta función proporciona a los usuarios que no son motores de ejecución de confianza credenciales para interactuar con los AWS servicios, si es necesario. No utilice este rol de IAM para permitir el acceso a los datos de Amazon S3, a menos que se trate de datos a los que deberían poder acceder todos los usuarios.

Esta función la asumirá el rol de perfil de EC2 instancia. Utilice el siguiente ejemplo para crear una política de confianza que permita que el rol de perfil de EC2 instancia asuma el rol de IAM de Apache Ranger.

```
    {
      "Sid": "",
      "Effect": "Allow",
      "Principal": {
        "AWS": "arn:aws:iam::<AWS_ACCOUNT_ID>:role/<EC2 INSTANCE PROFILE ROLE NAME eg.
EMR_EC2_DefaultRole>"
      },
      "Action": ["sts:AssumeRole", "sts:TagSession"]
```

```
    }
```

Valide sus permisos para la integración de Amazon EMR con Apache Ranger

Consulte Solución de problemas con Apache Ranger para obtener instrucciones sobre cómo validar los permisos.

Creación de la configuración de seguridad de EMR

Creación de una configuración de seguridad de Amazon EMR para Apache Ranger

Antes de lanzar un clúster de Amazon EMR integrado con Apache Ranger, cree una configuración de seguridad.

Console

Para crear una configuración de seguridad que especifique la opción de integración con AWS Ranger

1. En la consola de Amazon EMR, seleccione Configuraciones de seguridad y, a continuación, Crear.
2. Escriba un nombre para la configuración de seguridad en el campo Name (Nombre). Este nombre se utiliza para especificar la configuración de seguridad cuando crea un clúster.
3. En Integración con AWS Ranger, seleccione Habilitar un control de acceso detallado administrado por Apache Ranger.
4. Seleccione un rol de IAM para que Apache Ranger lo aplique. Para obtener más información, consulte Roles de IAM para la integración nativa con Apache Ranger.
5. Seleccione un rol de IAM para que otros servicios de AWS lo apliquen.
6. Configure los complementos para que se conecten al servidor Ranger Admin introduciendo el ARN de Secrets Manager del servidor de Admin y la dirección.
7. Seleccione las aplicaciones para configurar los complementos de Ranger. Introduzca el ARN de Secrets Manager que contiene el certificado TLS privado del complemento.

   Si no configura Apache Spark ni Apache Hive y los selecciona como aplicaciones para su clúster, la solicitud fallará.
8. Defina otras opciones de configuración de seguridad según corresponda y elija Create (Crear). Debe habilitar la autenticación de Kerberos mediante el KDC dedicado del clúster o un KDC externo.

> **Note**
>
> Actualmente, no puede utilizar la consola para crear una configuración de seguridad que especifique la opción de integración de AWS Ranger en el. AWS GovCloud (US) Region La configuración de seguridad se puede llevar a cabo con la CLI.

CLI

Para crear una configuración de seguridad para la integración de Apache Ranger

1. *<ACCOUNT ID>*Sustitúyala por tu ID AWS de cuenta.
2. Sustituya *<REGION>* por la región en la que se encuentra el recurso.
3. Especifique un valor para `TicketLifetimeInHours` para determinar el periodo para el que un vale de Kerberos generado por el KDC es válido.
4. Especifique la dirección del servidor de Ranger Admin de `AdminServerURL`.

```
{
    "AuthenticationConfiguration": {
        "KerberosConfiguration": {
            "Provider": "ClusterDedicatedKdc",
            "ClusterDedicatedKdcConfiguration": {
                "TicketLifetimeInHours": 24
            }
        }
    },
    "AuthorizationConfiguration":{
      "RangerConfiguration":{
         "AdminServerURL":"https://_<RANGER ADMIN SERVER IP>_:6182",
         "RoleForRangerPluginsARN":"arn:aws:iam::_<ACCOUNT ID>_:role/_<RANGER PLUGIN
 DATA ACCESS ROLE NAME>_",
         "RoleForOtherAWSServicesARN":"arn:aws:iam::_<ACCOUNT ID>_:role/_<USER
 ACCESS ROLE NAME>_",
         "AdminServerSecretARN":"arn:aws:secretsmanager:_<REGION>_:_<ACCOUNT
 ID>_:secret:_<SECRET NAME THAT PROVIDES ADMIN SERVERS PUBLIC TLS CERTIFICATE
 WITHOUT VERSION>_",
         "RangerPluginConfigurations":[
            {
               "App":"Spark",
```

```
                "ClientSecretARN":"arn:aws:secretsmanager:_<REGION>_:_<ACCOUNT
 ID>_:secret:_<SECRET NAME THAT PROVIDES SPARK PLUGIN PRIVATE TLS CERTIFICATE
 WITHOUT VERSION>_",
                "PolicyRepositoryName":"<SPARK SERVICE NAME eg. amazon-emr-spark>"
            },
            {
                "App":"Hive",
                "ClientSecretARN":"arn:aws:secretsmanager:_<REGION>_:_<ACCOUNT
 ID>_:secret:_<SECRET NAME THAT PROVIDES Hive PLUGIN PRIVATE TLS CERTIFICATE WITHOUT
 VERSION>_",
                "PolicyRepositoryName":"<HIVE SERVICE NAME eg. Hivedev>"
            },
            {
                "App":"EMRFS-S3",
                "ClientSecretARN":"arn:aws:secretsmanager:_<REGION>_:_<ACCOUNT
 ID>_:secret:_<SECRET NAME THAT PROVIDES EMRFS S3 PLUGIN PRIVATE TLS CERTIFICATE
 WITHOUT VERSION>_",
                "PolicyRepositoryName":"<EMRFS S3 SERVICE NAME eg amazon-emr-emrfs>"
            },
        {
                "App":"Trino",
                "ClientSecretARN":"arn:aws:secretsmanager:_<REGION>_:_<ACCOUNT
 ID>_:secret:_<SECRET NAME THAT PROVIDES TRINO PLUGIN PRIVATE TLS CERTIFICATE
 WITHOUT VERSION>_",
                "PolicyRepositoryName":"<TRINO SERVICE NAME eg amazon-emr-trino>"
            }
         ],
         "AuditConfiguration":{
            "Destinations":{
                "AmazonCloudWatchLogs":{
                   "CloudWatchLogGroup":"arn:aws:logs:<REGION>:_<ACCOUNT ID>_:log-
group:_<LOG GROUP NAME FOR AUDIT EVENTS>_"
                }
            }
         }
      }
   }
}
```

Estos PolicyRespositoryNames son los nombres de los servicios que se especifican en su administrador de Apache Ranger.

Cree una configuración de seguridad de Amazon EMR con el siguiente contenido. Sustituya la configuración de seguridad por el nombre de su elección. Seleccione esta configuración por el nombre al crear el clúster.

```
aws emr create-security-configuration \
--security-configuration file://./security-configuration.json \
--name security-configuration
```

Configuración de características de seguridad adicionales

Para integrar Amazon EMR de forma segura con Apache Range, configure las siguientes características de seguridad de EMR:

- Habilite la autenticación de Kerberos mediante el KDC dedicado del clúster o un KDC externo. Para obtener instrucciones, consulte Uso de Kerberos para la autenticación con Amazon EMR.
- (Opcional) Habilite el cifrado en tránsito o en reposo. Para obtener más información, consulte Opciones de cifrado para Amazon EMR.

Para obtener más información, consulte Seguridad en Amazon EMR.

Almacenamiento de certificados TLS en AWS Secrets Manager

Los complementos de Ranger instalados en un clúster de Amazon EMR y el servidor de Ranger Admin deben comunicarse a través del protocolo TLS para garantizar que los datos de la política y otra información enviada no se puedan leer si se interceptan. EMR también exige que los complementos se autentiquen en el servidor de Ranger Admin proporcionando su propio certificado TLS y que realicen una autenticación TLS bidireccional. Esta configuración requería la creación de cuatro certificados: dos pares de certificados TLS públicos y privados. Para obtener instrucciones sobre cómo instalar el certificado en su servidor de Ranger Admin, consulte Configure un servidor Ranger Admin para integrarlo con Amazon EMR. Para completar la configuración, los complementos de Ranger instalados en el clúster de EMR necesitan dos certificados: el certificado TLS público de su servidor de administración y el certificado privado que el complemento utilizará para autenticarse en el servidor de Ranger Admin. Para proporcionar estos certificados TLS, deben estar en una configuración de seguridad de EMR AWS Secrets Manager y proporcionarse en ella.

> **Note**
>
> Se recomienda encarecidamente, pero no es obligatorio, crear un par de certificados para cada una de sus aplicaciones a fin de limitar el impacto en caso de que uno de los certificados del complemento se vea comprometido.

> **Note**
>
> Debe realizar un seguimiento de los certificados y rotarlos antes de su fecha de vencimiento.

Formato del certificado

La importación de los certificados a la AWS Secrets Manager es la misma, independientemente de si se trata del certificado de complemento privado o del certificado de administrador de Ranger público. Antes de importar los certificados TLS, los certificados deben estar en formato PEM 509x.

Un ejemplo de certificado público tiene el siguiente formato:

```
-----BEGIN CERTIFICATE-----
...Certificate Body...
-----END CERTIFICATE-----
```

Un ejemplo de certificado privado tiene el siguiente formato:

```
-----BEGIN PRIVATE KEY-----
...Private Certificate Body...
-----END PRIVATE KEY-----
-----BEGIN CERTIFICATE-----
...Trust Certificate Body...
-----END CERTIFICATE-----
```

El certificado privado también debe contener un certificado de confianza.

Si desea obtener un certificado, debe ejecutar el siguiente comando:

```
openssl x509 -in <PEM FILE> -text
```

Importación de un certificado a AWS Secrets Manager

Al crear su secreto en Secrets Manager, seleccione Otro tipo de secretos en Tipo de secreto y pegue su certificado cifrado en PEM en el campo Texto no cifrado.

Inicio de un clúster de EMR con Apache Ranger

Antes de lanzar un clúster de Amazon EMR con Apache Ranger, asegúrese de que cada componente cumpla los siguientes requisitos mínimos de versión:

- Amazon EMR 5.32.0 o posterior, o 6.3.0 o posterior. Le recomendamos que utilice la última versión de Amazon EMR.
- Servidor de Apache Ranger Admin 2.x.

Siga estos pasos:

- Instale Apache Ranger si aún no lo ha hecho. Para obtener más información acerca de la instalación, consulte Apache Ranger 0.5.0 installation.
- Asegúrese de que haya conectividad de red entre el clúster de Amazon EMR y el servidor de Apache Ranger Admin. Consulte Configure un servidor Ranger Admin para integrarlo con Amazon EMR
- Cree los roles de IAM necesarios. Consulte Roles de IAM para la integración nativa con Apache Ranger.
- Cree una configuración de seguridad de EMR para la instalación de Apache Ranger. Para obtener más información, consulte Creación de la configuración de seguridad de EMR.

Configuración de Zeppelin para clústeres de Amazon EMR habilitados para Apache Ranger

El tema trata sobre cómo configurar Apache Zeppelin para un clúster de Amazon EMR compatible con Apache Ranger, de modo que pueda utilizar Zeppelin como cuaderno para navegar por los datos de forma interactiva. Zeppelin se incluye en las versiones de 5.0.0 de Amazon EMR y posteriores. Las versiones anteriores incluyen Zeppelin como una aplicación de entorno aislado. Para obtener más información, consulte Versiones de lanzamiento de Amazon EMR 4.x en la Guía de publicación de Amazon EMR.

De forma predeterminada, Zeppelin está configurado con una contraseña y un nombre de usuario predeterminados, lo que no es seguro en un entorno multiusuario.

Para configurar Zeppelin, siga los pasos que se describen a continuación:

1. Modifique el mecanismo de autenticación.

   Modifique el archivo `shiro.ini` para implementar el mecanismo de autenticación que prefiera. Zeppelin admite Active Directory, LDAP, PAM y Knox SSO. Consulte Apache Shiro authentication for Apache Zeppelin para obtener más información.

2. Configure Zeppelin para suplantar al usuario final

   Si permite que Zeppelin suplante al usuario final, los trabajos enviados por Zeppelin se pueden ejecutar como ese usuario final. Agregue la siguiente configuración a `core-site.xml`:

```
[
  {
    "Classification": "core-site",
    "Properties": {
      "hadoop.proxyuser.zeppelin.hosts": "*",
      "hadoop.proxyuser.zeppelin.groups": "*"
    },
    "Configurations": [
    ]
  }
]
```

   A continuación, agregue la siguiente configuración a `hadoop-kms-site.xml`, que se encuentra en `/etc/hadoop/conf`:

```
[
  {
    "Classification": "hadoop-kms-site",
    "Properties": {
      "hadoop.kms.proxyuser.zeppelin.hosts": "*",
      "hadoop.kms.proxyuser.zeppelin.groups": "*"
    },
    "Configurations": [
    ]
  }
]
```

También puede agregar estas configuraciones a su clúster de Amazon EMR mediante la consola siguiendo los pasos que se indican en Reconfiguración de un grupo de instancias en la consola.

3. Permita que Zeppelin utilice el comando sudo como usuario final

   Cree un archivo `/etc/sudoers.d/90-zeppelin-user` que contenga lo siguiente:

```
zeppelin ALL=(ALL) NOPASSWD:ALL
```

4. Modifique la configuración de los intérpretes para ejecutar las tareas de los usuarios en sus propios procesos.

   Configure todos los intérpretes para que creen instancias de los intérpretes “por usuario” en procesos “aislados”.

5. Modifique **`zeppelin-env.sh`**

   Agregue lo siguiente a `zeppelin-env.sh` para que Zeppelin comience a lanzar los intérpretes como usuario final:

```
ZEPPELIN_IMPERSONATE_USER=`echo ${ZEPPELIN_IMPERSONATE_USER} | cut -d @ -f1`
export ZEPPELIN_IMPERSONATE_CMD='sudo -H -u ${ZEPPELIN_IMPERSONATE_USER} bash -c'
```

   Agregue lo siguiente a `zeppelin-env.sh` para cambiar los permisos predeterminados del cuaderno a solo lectura únicamente para el creador:

```
export ZEPPELIN_NOTEBOOK_PUBLIC="false"
```

   Por último, añada lo siguiente `zeppelin-env.sh` para incluir la ruta de la RecordServer clase EMR después de la primera `CLASSPATH` sentencia:

```
export CLASSPATH="$CLASSPATH:/usr/share/aws/emr/record-server/lib/aws-emr-record-
server-connector-common.jar:/usr/share/aws/emr/record-server/lib/aws-emr-record-
server-spark-connector.jar:/usr/share/aws/emr/record-server/lib/aws-emr-record-
server-client.jar:/usr/share/aws/emr/record-server/lib/aws-emr-record-server-
common.jar:/usr/share/aws/emr/record-server/lib/jars/secret-agent-interface.jar"
```

6. Reinicie Zeppelin.

   Ejecute el siguiente comando para reiniciar Zeppelin:

```
sudo systemctl restart zeppelin
```

Problemas conocidos para la integración de Amazon EMR

Problemas conocidos

Existe un problema conocido en la versión 5.32 de Amazon EMR en el que se modificaron los permisos de `hive-site.xml` para que solo los usuarios con privilegios pudieran leerlo, ya que es posible que haya credenciales almacenadas en él. Esto podría impedir que Hue leyera `hive-site.xml` y provocar que las páginas web se recargaran continuamente. Si experimenta este problema, agregue la siguiente configuración para solucionarlo:

```
[
  {
    "Classification": "hue-ini",
    "Properties": {},
    "Configurations": [
      {
        "Classification": "desktop",
        "Properties": {
          "server_group":"hive_site_reader"
         },
        "Configurations":[
        ]
      }
    ]
  }
]
```

Existe un problema conocido que hace que el complemento EMRFS S3 para Apache Ranger no sea compatible actualmente con la característica Zona de seguridad de Apache Ranger. Las restricciones de control de acceso definidas mediante la característica Zona de seguridad no se aplican a los clústeres de Amazon EMR.

Aplicación UIs

De forma predeterminada, las IU de las aplicaciones no realizan la autenticación. Esto incluye la ResourceManager interfaz de usuario, la interfaz de NodeManager usuario y la interfaz de usuario

de Livy, entre otras. Además, cualquier usuario que tenga la posibilidad de acceder al UIs puede ver información sobre los trabajos de todos los demás usuarios.

Si no desea este comportamiento, debe asegurarse de que se utilice un grupo de seguridad para restringir el acceso de los usuarios a UIs la aplicación.

Permisos predeterminados de HDFS

De forma predeterminada, los objetos que los usuarios crean en HDFS reciben permisos legibles en todo el mundo. Esto puede provocar que los usuarios que no deberían tener acceder a los datos puedan leerlos. Para cambiar este comportamiento de forma que los permisos de archivo predeterminados sean de lectura y escritura únicamente para el creador del trabajo, lleve a cabo estos pasos.

Al crear el clúster de EMR, proporcione la siguiente configuración:

```
[
  {
    "Classification": "hdfs-site",
    "Properties": {
      "dfs.namenode.acls.enabled": "true",
      "fs.permissions.umask-mode": "077",
      "dfs.permissions.superusergroup": "hdfsadmingroup"
    }
  }
]
```

Además, ejecute la siguiente acción de arranque:

```
--bootstrap-actions Name='HDFS UMask Setup',Path=s3://elasticmapreduce/hdfs/umask/
umask-main.sh
```

Complementos Apache Ranger para escenarios de integración de Amazon EMR

Los complementos de Apache Ranger validan el acceso de un usuario según las políticas de autorización definidas en el servidor de administración de políticas de Apache Ranger.

Temas

- Complemento de Apache Hive para la integración de Ranger con Amazon EMR

- Complemento de Apache Spark para la integración de Ranger con Amazon EMR
- Complemento EMRFS S3 para la integración de Ranger con Amazon EMR
- Complemento Trino para la integración de Ranger con Amazon EMR

Complemento de Apache Hive para la integración de Ranger con Amazon EMR

Apache Hive es un popular motor de ejecución dentro del ecosistema Hadoop. Amazon EMR proporciona un complemento de Apache Ranger para poder proporcionar controles de acceso detallados para Hive. El complemento es compatible con la versión 2.0 y posteriores del servidor de Apache Ranger Admin de código abierto.

Temas

- Características admitidas
- Instalación de la configuración del servicio
- Consideraciones
- Limitaciones

Características admitidas

El complemento Apache Ranger para Hive en EMR admite todas las funciones del complemento de código abierto, que incluye controles de acceso a bases de datos, tablas y columnas, así como el filtrado de filas y el enmascaramiento de datos. Para ver una tabla de comandos de Hive y los permisos de Ranger asociados, consulte Hive commands to Ranger permission mapping.

Instalación de la configuración del servicio

El complemento Apache Hive es compatible con la definición de servicio de Hive existente en Apache Hive Hadoop SQL.

Si no tiene una instancia del servicio en Hadoop SQL, como se muestra arriba, puede crear una. Haga clic en el signo + situado junto a Hadoop SQL.

1. Nombre del servicio (si se muestra): ingrese el nombre del servicio. El valor sugerido es **`amazonemrhive`**. Anote el nombre de este servicio: es necesario al crear una configuración de seguridad de EMR.

2. Nombre público: ingrese el nombre que se mostrará para el servicio. El valor sugerido es **`amazonemrhive`**.

Las propiedades de Apache Hive Config se utilizan para establecer una conexión con su servidor Apache Ranger Admin con un HiveServer 2 para implementar el autocompletado al crear políticas. No es necesario que las siguientes propiedades sean precisas si no tiene un proceso persistente de HiveServer 2 y se pueden rellenar con cualquier información.

- Nombre de usuario: introduzca un nombre de usuario para la conexión JDBC a una instancia de HiveServer 2 instancias.
- Contraseña: ingrese la contraseña del nombre de usuario anterior.
- jdbc.driver. ClassName: Introduzca el nombre de la clase JDBC para la conectividad con Apache Hive. Puede utilizar el valor predeterminado.
- jdbc.url: Introduzca la cadena de conexión JDBC que se utilizará al conectarse a 2. HiveServer
- Nombre común del certificado: el campo CN (Nombre común) del certificado que se utiliza para conectarse al servidor de administración desde un complemento cliente. Este valor debe coincidir con el campo CN del certificado TLS que se creó para el complemento.

El botón Probar conexión comprueba si los valores anteriores se pueden utilizar para conectarse correctamente a la instancia 2. HiveServer Una vez que el servicio se haya creado correctamente, el administrador de servicios debería tener el siguiente aspecto:

Consideraciones

Servidor de metadatos de Hive

Solo motores fiables, específicamente Hive y `emr_record_server`, pueden acceder al servidor de metadatos de Hive para proteger al usuario del acceso no autorizado. Todos los nodos del clúster también acceden al servidor de metadatos de Hive. El puerto 9083 requerido proporciona a todos los nodos acceso al nodo principal.

Autenticación

De forma predeterminada, Apache Hive está configurado para autenticarse mediante Kerberos tal como se configuró en la configuración de seguridad de EMR. HiveServer2 también se puede configurar para autenticar a los usuarios mediante LDAP. Consulte Implementing LDAP authentication for Hive on a multi-tenant Amazon EMR cluster para obtener más información.

Limitaciones

Las siguientes son las limitaciones actuales del complemento Apache Hive en Amazon EMR 5.x:

- No se admiten roles de Hive de momento. No se admiten las instrucciones Grant ni Revoke.
- No se admite Hive CLI. JDBC/Beeline es la única forma autorizada de conectar Hive.
- `hive.server2.builtin.udf.blacklist`la configuración debe rellenarse con lo UDFs que considere inseguro.

Complemento de Apache Spark para la integración de Ranger con Amazon EMR

Amazon EMR ha integrado el EMR RecordServer para proporcionar un control de acceso detallado para SparkSQL. El EMR es un proceso privilegiado que RecordServer se ejecuta en todos los nodos de un clúster habilitado para Apache Ranger. Cuando un controlador o ejecutor de Spark ejecuta una sentencia de SparkSQL, todas las solicitudes de metadatos y datos pasan por. RecordServer Para obtener más información sobre EMR RecordServer, consulte la Componentes de Amazon EMR para su uso con Apache Ranger página.

Temas

- Características admitidas
- Reimplementación de la definición de servicio para usar las instrucciones INSERT, ALTER o DDL
- Instalación de la definición de servicio
- Creación de políticas de Spark SQL
- Consideraciones
- Limitaciones

Características admitidas

Acción SQL statement/Ranger	STATUS	Versión de EMR compatible
SELECT	Compatible	A partir de la 5.32

Acción SQL statement/Ranger	STATUS	Versión de EMR compatible
SHOW DATABASES	Compatible	A partir de la 5.32
SHOW COLUMNS	Compatible	A partir de la 5.32
SHOW TABLES	Compatible	A partir de la 5.32
MOSTRAR LAS PROPIEDADES DE LA TABLA	Compatible	A partir de la 5.32
DESCRIBE TABLE	Compatible	A partir de la 5.32
INSERT OVERWRITE	Compatible	A partir de las 5.34 y 6.4
INSERT INTO	Compatible	A partir de las 5.34 y 6.4
ALTER TABLE	Compatible	A partir de la 6.4
CREATE TABLE	Compatible	A partir de las 5.35 y 6.7
CREATE DATABASE	Compatible	A partir de las 5.35 y 6.7
DROP TABLE	Compatible	A partir de las 5.35 y 6.7
DROP DATABASE	Compatible	A partir de las 5.35 y 6.7
DROP VIEW	Compatible	A partir de las 5.35 y 6.7
CREATE VIEW	No es compatible	

Se admiten las siguientes características al usar Spark SQL:

- Control de acceso detallado a las tablas del metaalmacén de Hive, y las políticas se pueden crear para bases de datos, tablas y columnas.
- Las políticas de Apache Ranger pueden incluir políticas de concesión y de denegación a usuarios y grupos.
- Los eventos de auditoría se envían a CloudWatch los registros.

Reimplementación de la definición de servicio para usar las instrucciones INSERT, ALTER o DDL

> **Note**
>
> A partir de Amazon EMR 6.4, puede usar Spark SQL con las instrucciones: INSERT INTO, INSERT OVERWRITE o ALTER TABLE. A partir de Amazon EMR 6.7, puede usar Spark SQL para crear o eliminar bases de datos y tablas. Si ya tiene una instalación en el servidor de Apache Ranger con las definiciones de servicio de Apache Spark implementadas, utilice el siguiente código para volver a implementar las definiciones de servicio.

```
# Get existing Spark service definition id calling Ranger REST API and JSON
 processor
curl --silent -f -u <admin_user_login>:<password_for_ranger_admin_user> \
-H "Accept: application/json" \
-H "Content-Type: application/json" \
-k 'https://*<RANGER SERVER ADDRESS>*:6182/service/public/v2/api/servicedef/
name/amazon-emr-spark' | jq .id

# Download the latest Service definition
wget https://s3.amazonaws.com/elasticmapreduce/ranger/service-definitions/
version-2.0/ranger-servicedef-amazon-emr-spark.json

# Update the service definition using the Ranger REST API
curl -u <admin_user_login>:<password_for_ranger_admin_user> -X PUT -d @ranger-
servicedef-amazon-emr-spark.json \
-H "Accept: application/json" \
-H "Content-Type: application/json" \
-k 'https://*<RANGER SERVER ADDRESS>*:6182/service/public/v2/api/
servicedef/<Spark service definition id from step 1>'
```

Instalación de la definición de servicio

La instalación de la definición de servicio Apache Spark de EMR requiere la configuración del servidor de Ranger Admin. Consulte Configure un servidor Ranger Admin para integrarlo con Amazon EMR.

Siga estos pasos para instalar la definición de servicio de Apache Spark:

Paso 1: inicie sesión mediante SSH en el servidor de Apache Ranger Admin

Por ejemplo:

```
ssh ec2-user@ip-xxx-xxx-xxx-xxx.ec2.internal
```

Paso 2: descargue la definición de servicio y el complemento del servidor de Apache Ranger Admin

En un directorio temporal, descargue la definición de servicio. Esta definición de servicio es compatible con las versiones 2.x de Ranger.

```
mkdir /tmp/emr-spark-plugin/
cd /tmp/emr-spark-plugin/

wget https://s3.amazonaws.com/elasticmapreduce/ranger/service-definitions/version-2.0/
ranger-spark-plugin-2.x.jar
wget https://s3.amazonaws.com/elasticmapreduce/ranger/service-definitions/version-2.0/
ranger-servicedef-amazon-emr-spark.json
```

Paso 3: instale el complemento Apache Spark para Amazon EMR

```
export RANGER_HOME=.. # Replace this Ranger Admin's home directory eg /usr/lib/ranger/
ranger-2.0.0-admin
mkdir $RANGER_HOME/ews/webapp/WEB-INF/classes/ranger-plugins/amazon-emr-spark
mv ranger-spark-plugin-2.x.jar $RANGER_HOME/ews/webapp/WEB-INF/classes/ranger-plugins/
amazon-emr-spark
```

Paso 4: registre la definición de servicio de Apache Spark para Amazon EMR

```
curl -u *<admin users login>*:*_<_**_password_ **_for_** _ranger admin user_**_>_* -X
 POST -d @ranger-servicedef-amazon-emr-spark.json \
-H "Accept: application/json" \
```

```
-H "Content-Type: application/json" \
-k 'https://*<RANGER SERVER ADDRESS>*:6182/service/public/v2/api/servicedef'
```

Si este comando se ejecuta correctamente, verá un nuevo servicio en la IU de Ranger Admin denominado “AMAZON-EMR-SPARK”, como se muestra en la siguiente imagen (se muestra la versión 2.0 de Ranger).

Paso 5: Crear una instancia de la AMAZON-EMR-SPARK aplicación

Nombre del servicio (si se muestra): el nombre del servicio que se utilizará. El valor sugerido es **`amazonemrspark`**. Anote el nombre de este servicio, ya que será necesario al crear una configuración de seguridad de EMR.

Nombre público: el nombre que se mostrará para esta instancia. El valor sugerido es **`amazonemrspark`**.

Nombre común del certificado: el campo CN (Nombre común) del certificado que se utiliza para conectarse al servidor de administración desde un complemento cliente. Este valor debe coincidir con el campo CN del certificado TLS que se creó para el complemento.

> **Note**
>
> El certificado TLS de este complemento debería haberse registrado en el almacén de confianza del servidor Ranger Admin. Consulte Certificados TLS para la integración de Apache Ranger con Amazon EMR para obtener más detalles.

Creación de políticas de Spark SQL

Al crear una nueva política, los campos que hay que rellenar son:

Nombre de la política: el nombre de la política.

Etiqueta de la política: una etiqueta que puede poner en esta política.

Base de datos: la base de datos a la que se aplica esta política. El comodín “*” representa todas las bases de datos.

Tabla: las tablas a las que se aplica esta política. El comodín “*” representa todas las tablas.

Columna de EMR Spark: las columnas a las que se aplica esta política. El comodín “*” representa todas las columnas.

Descripción: una descripción de esta política.

Para especificar los usuarios y grupos, ingrese los usuarios y grupos que aparecen a continuación para conceder los permisos. También puede especificar exclusiones para las condiciones de autorización y denegación.

Tras especificar las condiciones de autorización y denegación, haga clic en Guardar.

Consideraciones

Cada nodo del clúster de EMR debe poder conectarse al nodo principal en el puerto 9083.

Limitaciones

Las siguientes son las limitaciones actuales del complemento Apache Spark:

- Servidor de registros siempre se conectará al HMS que se ejecute en un clúster de Amazon EMR. Configure el HMS para que se conecte al modo remoto, si es necesario. No debe incluir valores de configuración en el archivo de configuración Hive-site.xml de Apache Spark.
- Las tablas creadas con fuentes de datos de Spark en CSV o Avro no se pueden leer con EMR. RecordServer Use Hive para crear y escribir datos, y lea con Record.
- No se admiten las tablas Delta Lake, Hudi e Iceberg.
- Los usuarios tienen que tener acceso a la base de datos predeterminada. Este es un requisito para el uso de Apache Spark.
- El servidor de Ranger Admin no admite la característica de autocompletar.
- El complemento SparkSQL para Amazon EMR no admite filtros de filas ni enmascaramiento de datos.
- Al utilizar ALTER TABLE con Spark SQL, la ubicación de una partición debe ser el directorio secundario de la ubicación de una tabla. No se admite la inserción de datos en una partición cuya ubicación sea diferente de la ubicación de la tabla.

Complemento EMRFS S3 para la integración de Ranger con Amazon EMR

Para facilitar los controles de acceso a los objetos de S3 en un clúster multiusuario, el complemento EMRFS S3 proporciona controles de acceso a los datos de S3 cuando se accede a ellos a través de EMRFS. Puede permitir el acceso a los recursos de usuarios y grupos de S3.

Para ello, cuando su aplicación intente acceder a los datos de S3, EMRFS envía una solicitud de credenciales al proceso del agente secreto, donde la solicitud se autentica y autoriza mediante un complemento de Apache Ranger. Si la solicitud se autoriza, el agente secreto asume el rol de IAM para los motores Apache Ranger con una política restringida para generar credenciales que solo tienen acceso a la política de Ranger que permitió el acceso. A continuación, las credenciales se devuelven a EMRFS para acceder a S3.

Temas

- Características admitidas
- Instalación de la configuración del servicio
- Creación de políticas de EMRFS S3
- Notas de uso de las políticas de EMRFS S3
- Limitaciones

Características admitidas

El complemento EMRFS S3 proporciona autorización de almacenamiento. Se pueden crear políticas para proporcionar acceso a los usuarios y grupos a los buckets y prefijos de S3. La autorización se realiza únicamente para EMRFS.

Instalación de la configuración del servicio

Para instalar la definición del servicio EMRFS, debe configurar el servidor Ranger Admin. Para configurar el servidor, consulte Configure un servidor Ranger Admin para integrarlo con Amazon EMR.

Siga estos pasos para instalar la definición de servicio de EMRFS.

Paso 1: inicie sesión mediante SSH en el servidor de Apache Ranger Admin.

Por ejemplo:

```
ssh ec2-user@ip-xxx-xxx-xxx-xxx.ec2.internal
```

Paso 2: descargue la definición del servicio de EMRFS.

En un directorio temporal, descargue la definición de servicio de Amazon EMR. Esta definición de servicio es compatible con las versiones 2.x de Ranger.

```
wget https://s3.amazonaws.com/elasticmapreduce/ranger/service-definitions/version-2.0/
ranger-servicedef-amazon-emr-emrfs.json
```

Paso 3: registre la definición de servicio de EMRFS S3.

```
curl -u *<admin users login>*:*_<_**_password_ **_for_** _ranger admin user_**_>_* -X
 POST -d @ranger-servicedef-amazon-emr-emrfs.json \
-H "Accept: application/json" \
-H "Content-Type: application/json" \
-k 'https://*<RANGER SERVER ADDRESS>*:6182/service/public/v2/api/servicedef'
```

Si este comando se ejecuta correctamente, verá un nuevo servicio en la IU de Ranger Admin denominado “AMAZON-EMR-S3”, como se muestra en la siguiente imagen (se muestra la versión 2.0 de Ranger).

Paso 4: Crea una instancia de la AMAZON-EMR-EMRFS aplicación.

Cree una instancia de la definición de servicio.

- Haga clic en el signo + situado junto a AMAZON-EMR-EMRFS.

Rellene los siguientes campos:

Nombre del servicio (si se muestra): el valor sugerido es **`amazonemrs3`**. Anote el nombre de este servicio, ya que será necesario al crear una configuración de seguridad de EMR.

Nombre público: el nombre que se muestra para este servicio. El valor sugerido es **`amazonemrs3`**.

Nombre común del certificado: el campo CN (Nombre común) del certificado que se utiliza para conectarse al servidor de administración desde un complemento cliente. Este valor debe coincidir con el campo CN del certificado TLS que se creó para el complemento.

> **Note**
>
> El certificado TLS de este complemento debería haberse registrado en el almacén de confianza del servidor Ranger Admin. Consulte Certificados TLS para la integración de Apache Ranger con Amazon EMR para obtener más detalles.

Cuando se crea el servicio, el administrador de servicios incluye “AMAZON-EMR-EMRFS”, como se muestra en la siguiente imagen.

Creación de políticas de EMRFS S3

Para crear una nueva política en la página Crear política del administrador de servicios, rellene los siguientes campos.

Nombre de la política: el nombre de la política.

Etiqueta de la política: una etiqueta que puede poner en esta política.

Recurso de S3: un recurso que comienza con el bucket y el prefijo opcional. Consulte Notas de uso de las políticas de EMRFS S3 para obtener más información sobre las prácticas recomendadas. Los recursos del servidor Ranger Admin no deben contener **`s3://`**, **`s3a://`** ni **`s3n://`**.

Puede especificar los usuarios y grupos a los que conceder permisos. También puede especificar exclusiones para las condiciones de autorización y denegación.

> **Note**
>
> Se permite un máximo de tres recursos para cada política. Agregar más de tres recursos puede provocar un error cuando se usa esta política en un clúster de EMR. Al agregar más de tres políticas, se muestra un recordatorio sobre el límite de la política.

Notas de uso de las políticas de EMRFS S3

Al crear políticas de S3 en Apache Ranger, hay que tener en cuenta algunas consideraciones de uso.

Permisos para varios objetos de S3

Puede utilizar políticas recursivas y expresiones comodín para conceder permisos a varios objetos de S3 con prefijos comunes. Las políticas recursivas otorgan permisos a todos los objetos con un prefijo común. Las expresiones comodín seleccionan varios prefijos. En conjunto, otorgan permisos a todos los objetos con varios prefijos comunes, como se muestra en los ejemplos siguientes.

Example Uso de una política recursiva

Supongamos que desea obtener permisos para enumerar todos los archivos Parquet de un bucket de S3 organizados de la siguiente manera.

```
s3://sales-reports/americas/
    +- year=2000
    |      +- data-q1.parquet
    |      +- data-q2.parquet
    +- year=2019
    |      +- data-q1.json
    |      +- data-q2.json
    |      +- data-q3.json
    |      +- data-q4.json
    |
    +- year=2020
    |      +- data-q1.parquet
    |      +- data-q2.parquet
    |      +- data-q3.parquet
    |      +- data-q4.parquet
    |      +- annual-summary.parquet
    +- year=2021
```

En primer lugar, considere los archivos Parquet con el prefijo `s3://sales-reports/americas/year=2000`. Puede conceder GetObject permisos a todos ellos de dos maneras:

Uso de políticas no recursivas: una opción es utilizar dos políticas no recursivas independientes, una para el directorio y otra para los archivos.

La primera política concede permiso al prefijo `s3://sales-reports/americas/year=2020` (no hay ningún `/` final).

```
- S3 resource = "sales-reports/americas/year=2000"
- permission = "GetObject"
- user = "analyst"
```

La segunda política utiliza una expresión comodín para conceder permisos a todos los archivos con prefijo `sales-reports/americas/year=2020/` (tenga en cuenta el `/` final).

```
- S3 resource = "sales-reports/americas/year=2020/*"
- permission = "GetObject"
- user = "analyst"
```

Uso de una política recursiva: una alternativa más práctica consiste en utilizar una única política recursiva y conceder permisos recursivos al prefijo.

```
 - S3 resource = "sales-reports/americas/year=2020"
 - permission = "GetObject"
 - user = "analyst"
 - is recursive = "True"
```

Hasta ahora, solo se han incluido los archivos Parquet con el prefijo `s3://sales-reports/americas/year=2000`. Ahora también puede incluir los archivos Parquet con un prefijo diferente, `s3://sales-reports/americas/year=2020`, en la misma política recursiva introduciendo una expresión comodín como la siguiente.

```
 - S3 resource = "sales-reports/americas/year=20?0"
 - permission = "GetObject"
 - user = "analyst"
 - is recursive = "True"
```

Políticas PutObject y DeleteObject permisos

La redacción de políticas `PutObject` y `DeleteObject` permisos para los archivos en EMRFS requiere un cuidado especial porque, a diferencia de GetObject los permisos, requieren la concesión de permisos recursivos adicionales al prefijo.

Example Políticas y permisos PutObject DeleteObject

Por ejemplo, para eliminar el archivo no solo se `annual-summary.parquet` requiere un DeleteObject permiso para acceder al archivo real.

```
- S3 resource = "sales-reports/americas/year=2020/annual-summary.parquet"
- permission = "DeleteObject"
- user = "analyst"
```

También requiere una política que conceda permisos `PutObject` y `GetObject` recursivos a su prefijo.

Del mismo modo, modificar el archivo `annual-summary.parquet` no solo requiere un permiso `PutObject` para ese archivo concreto.

```
- S3 resource = "sales-reports/americas/year=2020/annual-summary.parquet"
- permission = "PutObject"
- user = "analyst"
```

También requiere una política que conceda permisos `GetObject` recursivos a su prefijo.

```
- S3 resource = "sales-reports/americas/year=2020"
- permission = "GetObject"
- user = "analyst"
- is recursive = "True"
```

Comodines en las políticas

Hay dos áreas en las que se pueden especificar los caracteres comodín. Al especificar un recurso de S3, los caracteres “*” y “?” se puede utilizar. El asterisco “*” sustituye una ruta de S3 y coincide con todo lo que aparece después del prefijo. Tome la siguiente política como ejemplo:

```
S3 resource = "sales-reports/americas/*"
```

Esto coincide con las siguientes rutas de S3.

```
sales-reports/americas/year=2020/
sales-reports/americas/year=2019/
sales-reports/americas/year=2019/month=12/day=1/afile.parquet
sales-reports/americas/year=2018/month=6/day=1/afile.parquet
sales-reports/americas/year=2017/afile.parquet
```

El comodín “?” coincide con cualquier carácter individual. Tome la siguiente política como ejemplo:

```
S3 resource = "sales-reports/americas/year=201?/"
```

Esto coincide con las siguientes rutas de S3.

```
sales-reports/americas/year=2019/
```

```
sales-reports/americas/year=2018/
sales-reports/americas/year=2017/
```

Comodines en los usuarios

Hay dos caracteres comodín integrados al asignar usuarios para proporcionar acceso a los usuarios. El primero es el comodín “{USER}”, que proporciona acceso a todos los usuarios. El segundo comodín es “{OWNER}”, que proporciona acceso al propietario de un objeto concreto o directamente. Sin embargo, actualmente no se admite el comodín “{USER}”.

Limitaciones

Las siguientes son las limitaciones actuales del complemento EMRFS S3:

- Las políticas de Apache Ranger pueden tener como máximo tres políticas.
- El acceso a S3 debe realizarse a través de EMRFS y se puede utilizar con aplicaciones relacionadas con Hadoop. No se admite lo siguiente:

  – Bibliotecas Boto3

  - AWS SDK y CLI de AWK

  – Conector de código abierto S3A
- No se admiten las políticas de denegación de Apache Ranger.
- Actualmente, no se admiten las operaciones en S3 con claves con cifrado CSE-KMS.
- No se admite la compatibilidad entre regiones.
- La característica Zona de seguridad de Apache Ranger no es compatible. Las restricciones de control de acceso definidas mediante la característica Zona de seguridad no se aplican a los clústeres de Amazon EMR.
- El usuario de Hadoop no genera ningún evento de auditoría, ya que Hadoop siempre accede al perfil de la instancia. EC2
- Se recomienda deshabilitar la visión de consistencia de Amazon EMR. S3 de tiene un alto grado de consistencia, por lo que ya no es necesario. Consulte Amazon S3 strong consistency para obtener más información.
- El complemento EMRFS S3 realiza numerosas llamadas STS. Se recomienda realizar pruebas de carga en una cuenta de desarrollo y supervisar el volumen de llamadas STS. También se recomienda realizar una solicitud de STS para aumentar los límites del servicio. AssumeRole
- El servidor Ranger Admin no admite la característica de autocompletar.

Complemento Trino para la integración de Ranger con Amazon EMR

Trino (anteriormente PrestoSQL) es un motor de consultas SQL que puede utilizar para ejecutar consultas en orígenes de datos como HDFS, almacenamiento de objetos, bases de datos relacionales y bases de datos NoSQL. Elimina la necesidad de migrar los datos a una ubicación central y permite consultar estos datos desde cualquier lugar donde estén. Amazon EMR proporciona un complemento de Apache Ranger para proporcionar controles de acceso detallados para Trino. El complemento es compatible con la versión 2.0 y posteriores del servidor de Apache Ranger Admin de código abierto.

Temas

- Características admitidas
- Instalación de la configuración del servicio
- Creación de políticas de Trino
- Consideraciones
- Limitaciones

Características admitidas

El complemento Apache Ranger para Trino en Amazon EMR admite todas las funciones del motor de consultas de Trino, que están protegidas por un control de acceso detallado. Esto incluye controles de acceso a bases de datos, tablas y columnas, así como el filtrado de filas y el enmascaramiento de datos. Las políticas de Apache Ranger pueden incluir políticas de concesión y de denegación a usuarios y grupos. Los eventos de auditoría también se envían a CloudWatch los registros.

Instalación de la configuración del servicio

La instalación de la definición de servicio de Trino requiere la configuración del servidor de Ranger Admin. Para configurar el servidor de Ranger Admin, consulte Configure un servidor Ranger Admin para integrarlo con Amazon EMR.

Siga estos pasos para instalar la definición de servicio de Trino.

1. Inicie sesión mediante SSH en el servidor de Apache Ranger Admin.

```
ssh ec2-user@ip-xxx-xxx-xxx-xxx.ec2.internal
```

2. Desinstale el complemento del servidor Presto, si existe. Ejecute el siguiente comando. Si se produce un error “Servicio no encontrado”, significa que el complemento del servidor Presto no estaba instalado en su servidor. Continúe con el siguiente paso.

```
curl -f -u *<admin users login>*:*_<_**_password_ **_for_** _ranger admin
 user_**_>_* -X DELETE -k 'https://*<RANGER SERVER ADDRESS>*:6182/service/public/
v2/api/servicedef/name/presto'
```

3. Descargue la definición de servicio y el complemento del servidor de Apache Ranger Admin. En un directorio temporal, descargue la definición de servicio. Esta definición de servicio es compatible con las versiones 2.x de Ranger.

```
wget https://s3.amazonaws.com/elasticmapreduce/ranger/service-definitions/
version-2.0/ranger-servicedef-amazon-emr-trino.json
```

4. Registre la definición de servicio de Apache Trino para Amazon EMR.

```
curl -u *<admin users login>*:*_<_**_password_ **_for_** _ranger admin user_**_>_*
 -X POST -d @ranger-servicedef-amazon-emr-trino.json \
-H "Accept: application/json" \
-H "Content-Type: application/json" \
-k 'https://*<RANGER SERVER ADDRESS>*:6182/service/public/v2/api/servicedef'
```

Si este comando se ejecuta correctamente, verá un nuevo servicio en la IU de Ranger Admin denominado `TRINO`, como se muestra en la siguiente imagen.

5. Cree una instancia de la aplicación `TRINO` e ingrese la siguiente información.

Nombre del servicio: el nombre del servicio que utilizará. El valor sugerido es `amazonemrtrino`. Anote el nombre de este servicio, ya que será necesario al crear una configuración de seguridad de Amazon EMR.

Nombre público: el nombre que se mostrará para esta instancia. El valor sugerido es `amazonemrtrino`.

jdbc.driver. ClassName: El nombre de clase de la clase JDBC para la conectividad Trino. Puede utilizar el valor predeterminado.

jdbc.url: la cadena de conexión JDBC que se utilizará al conectarse al coordinador de Trino.

Nombre común del certificado: el campo CN (Nombre común) del certificado que se utiliza para conectarse al servidor de administración desde un complemento cliente. Este valor debe coincidir con el campo CN del certificado TLS que se creó para el complemento.

Tenga en cuenta que el certificado TLS de este complemento debería haberse registrado en el almacén de confianza del servidor de Ranger Admin. Para obtener más información, consulte Certificados TLS.

Creación de políticas de Trino

Cuando cree una nueva política, rellena los siguientes campos.

Nombre de la política: el nombre de la política.

Etiqueta de la política: una etiqueta que puede poner en esta política.

Catálogo: el catálogo al que se aplica esta política. El comodín “*” representa todos los catálogos.

Esquema: los esquemas a los que se aplica esta política. El comodín “*” representa todos los esquemas.

Tabla: las tablas a las que se aplica esta política. El comodín “*” representa todas las tablas.

Columna: las columnas a las que se aplica esta política. El comodín “*” representa todas las columnas.

Descripción: una descripción de esta política.

Existen otros tipos de políticas para Usuario de Trino (para el acceso suplantando al usuario), Propiedad del sistema o sesión de Trino (para modificar las propiedades del sistema o la sesión del motor), Funciones/Procedimientos (para permitir las llamadas a funciones o procedimientos) y URL (para conceder acceso de lectura/escritura al motor en las ubicaciones de los datos).

Para conceder permisos a usuarios y grupos específicos, ingrese los usuarios y grupos. También puede especificar exclusiones para las condiciones de autorización y denegación.

Tras especificar las condiciones de autorización y denegación, elija Guardar.

Consideraciones

Al crear políticas de Trino en Apache Ranger, hay que tener en cuenta algunas consideraciones de uso.

Servidor de metadatos de Hive

Solo motores fiables, específicamente el motor Trino, pueden acceder al servidor de metadatos de Hive para proteger al usuario del acceso no autorizado. Todos los nodos del clúster también acceden al servidor de metadatos de Hive. El puerto 9083 requerido proporciona a todos los nodos acceso al nodo principal.

Autenticación

De forma predeterminada, Trino está configurado para autenticarse mediante Kerberos, tal y como se indicó en la configuración de seguridad de Amazon EMR.

Cifrado en tránsito obligatorio

El complemento Trino requiere que tenga activado el cifrado en tránsito en la configuración de seguridad de Amazon EMR. Para habilitar el cifrado, consulte Cifrado en tránsito.

Limitaciones

Las siguientes son las limitaciones actuales del complemento Trino:

- El servidor de Ranger Admin no admite la característica de autocompletar.

Solución de problemas con Apache Ranger

Estos son algunos de los problemas que se diagnostican con frecuencia relacionados con el uso de Apache Ranger.

Recomendaciones

- Haga pruebas con un único clúster de nodo principal: los clústeres maestros de un solo nodo se aprovisionan más rápido que un clúster de varios nodos, lo que puede reducir el tiempo de cada iteración de prueba.
- Configure el modo de desarrollo del clúster. Al iniciar el clúster de EMR, establezca el parámetro `--additional-info"` en:

  ```
  '{"clusterType":"development"}'
  ```

Este parámetro solo se puede configurar mediante la AWS CLI o el AWS SDK y no está disponible a través de la consola Amazon EMR. Cuando se activa este indicador y el clúster maestro no lo aprovisiona, el servicio Amazon EMR mantiene activo el clúster durante algún tiempo antes de retirarlo del servicio. Este tiempo es muy útil para sondear varios archivos de registro antes de que finalice el clúster.

El clúster de EMR no se pudo aprovisionar

Existen varios motivos por los que un clúster de Amazon EMR puede no iniciarse. Las siguientes son algunas formas de diagnosticar el problema.

Compruebe los registros de aprovisionamiento de EMR

Amazon EMR usa Puppet para instalar y configurar aplicaciones en un clúster. Si consulta los registros, obtendrá detalles sobre si se ha producido algún error durante la fase de aprovisionamiento de un clúster. Se puede acceder a los registros en el clúster o en S3 si los registros están configurados para enviarse a S3.

Los registros se almacenan en `/var/log/provision-node/apps-phase/0/{UUID}/puppet.log` en el disco y `s3://<LOG LOCATION>/<CLUSTER ID>/node/<EC2 INSTANCE ID>/provision-node/apps-phase/0/{UUID}/puppet.log.gz`.

Mensajes de error comunes

Mensaje de error	Causa
Marioneta (error): ¡Error al iniciar el sistema! emr-record-server log de journalctl para: emr-record-server	No se pudo iniciar Servidor de registros de EMR. Consulte los registros de Servidor de registros de EMR a continuación.
Marioneta (error): ¡Error al iniciar el sistema! emr-record-server Registro journalctl para emrsecretagent:	Agente secreto de EMR no se pudo iniciar. Consulte Revisión de los registros de Agente secreto a continuación.
/Stage [main] /Ranger_Plugins: :Ranger_hive_ (aviso): 140408606197664:error:0906D	El certificado TLS privado de Secrets Manager para el certificado del complemento Apache

Mensaje de error	Causa
06C:PEM Rutines:PEM_READ_BIO:Sin línea de inicio:PEM_LIB.C:707:En esperaplugin/ Ranger_plugins::Prepare_two_way_tls[confi gure 2-way TLS in Hive plugin]/Exec[creat e keystore and truststore for Ranger Hive plugin]/returns: cualquier clave privada	Ranger no tiene el formato correcto o no es un certificado privado. Consulte Certificados TLS para la integración de Apache Ranger con Amazon EMR para obtener más información sobre los formatos de los certificados.
/Stage [main] /Ranger_Plugins: :ranger_S3_ plugin/Ranger_plugins::Prepare_two_way_tls[co nfigure 2-way TLS in Ranger s3 plugin]/E xec[create keystore and truststore for Ranger amazon-emr-s3 plugin]/returns (notice): An error occurred (AccessDeniedException) when calling the GetSecretValue operation: User: arn:aws:sts::XXXXXXXXXXX:assumed-role/EMR_EC2_DefaultRole/i -XXXXXXXX XX no está autorizado a actuar: secretsma nager: GetSecretValue on resource: arn:aws:s ecretsmanager:us-east-1:xxxxxxxxxx:secret: -XXXXX AdminServer	El EC2 rol del perfil de instancia no tiene los permisos correctos para recuperar los certifica dos TLS de Secrets Agent.

Compruebe los registros SecretAgent

Los registros de Agente secreto se encuentran en `/emr/secretagent/log/` en un nodo de EMR o en el directorio `s3:///<LOG LOCATION>/<CLUSTER ID>/node/<EC2 INSTANCE ID>/daemons/secretagent/` de S3.

Mensajes de error comunes

Mensaje de error	Causa
Excepción en el hilo «principal» com.amazo naws.services.securitytoken.model. AWSSecuri tyTokenServiceException: Usuario: arn:aws:s	La excepción anterior significa que el rol del perfil de la EC2 instancia EMR no tiene permisos para asumir el rol. RangerPlu

Mensaje de error	Causa
ts: :xxxxxxxxxxxx:Assumed- role/EMR_ EC2_DefaultRole/i -XXXXXXXXXXXXXXX no está autorizado a realizar: sts: AssumeRole on resource: arn:aws:iam: RangerPluginDataAc cessRole :xxxxxxxxxxxx:role/* (Servicio:; Código de estado: 403; Código de error:; ID de solicitud: XXXXXXXX-XXXXXX-XX XXXXXXXXXXXX AWSSecurityTokenService; Proxy: null) AccessDenied	ginDataAccessRole Consulte Roles de IAM para la integración nativa con Apache Ranger.
ERROR qtp54617902-149: Web App Exception Occurred javax.ws.rs. NotAllowedException: El método HTTP 405 no está permitido	Estos errores se pueden ignorar.

Compruebe los registros de Servidor de registros (para Spark SQL)

Los registros del servidor de registros EMR están disponibles en/var/log/emr-record-server/ en un nodo EMR, o se encuentran en el directorio s3:<LOG LOCATION>////node/ < INSTANCE <CLUSTER ID>EC2 ID>/Daemons//en S3. emr-record-server

Mensajes de error comunes

Mensaje de error	Causa
InstanceMetadataServiceResourceFetch erLos registros del servidor de registros de EMR están disponibles en/-record-server/ en un nodo de EMR, o en el directorio s3:///nod e/ < INSTANCE ID>/DAEMONS//de S3. ---- sep----:105 - [] No se pudo recuperar el token com.amazonaws. SdkClientException: No se pudo conectar al punto final del servicio	El EMR SecretAgent no apareció o está teniendo un problema. Inspeccione los SecretAgent registros para ver si hay errores y el script de la marioneta para determinar si hubo algún error de aprovisionamiento.

Las consultas fallan inesperadamente para la integración de Ranger con Amazon EMR

Compruebe los registros del complemento Apache Ranger (registros de Apache Hive, EMRRecordServer, SecretAgent EMR, etc.)

Esta sección es común a todas las aplicaciones que se integran con el complemento Ranger, como Apache Hive, EMR Record Server y EMR. SecretAgent

Mensajes de error comunes

Mensaje de error	Causa
ERROR:272 PolicyRefresher - [] (PolicyRefresherserviceName=policy-repository): no se pudo encontrar el servicio. Will clean up local cache of policies (-1)	Este mensaje de error significa que el nombre del servicio que proporcionó en la configuración de seguridad de EMR no coincide con un repositorio de políticas de servicio del servidor de Ranger Admin.

Si en el servidor Ranger Admin su AMAZON-EMR-SPARK servicio tiene el siguiente aspecto, debe ingresarlo como nombre del servicio. **`amazonemrspark`**

Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR

Los grupos de seguridad actúan como firewalls virtuales para las EC2 instancias del clúster a fin de controlar el tráfico entrante y saliente. Cada grupo de seguridad tiene un conjunto de reglas que controlan el tráfico entrante y un conjunto de reglas distinto que controlan el tráfico saliente. Para obtener más información, consulta los grupos EC2 de seguridad de Amazon para instancias de Linux en la Guía del EC2 usuario de Amazon.

Puede utilizar dos clases de grupos de seguridad con Amazon EMR: grupos de seguridad administrados por Amazon EMR y grupos de seguridad adicionales.

Cada clúster tiene asociados grupos de seguridad administrados. Puede utilizar los grupos de seguridad administrados predeterminados que Amazon EMR crea o especificar grupos de seguridad

administrados personalizados. De cualquier forma, Amazon EMR añade automáticamente las reglas a los grupos de seguridad gestionados que un clúster necesita para comunicarse entre las instancias y AWS los servicios del clúster.

Los grupos de seguridad adicionales son opcionales. Puede especificarlos junto con los grupos de seguridad administrados para adaptar el acceso a las instancias de clúster. Los grupos de seguridad adicionales contienen solo las reglas que defina. Amazon EMR no los modifica.

Las reglas que Amazon EMR crea en los grupos de seguridad administrados permiten la comunicación entre los componentes internos del clúster. Para permitir que los usuarios y las aplicaciones obtengan acceso a un clúster desde fuera de este, puede editar las reglas de los grupos de seguridad administrados, crear grupos de seguridad adicionales con reglas adicionales o ambas cosas.

> **Important**
>
> Además, la edición de reglas de los grupos de seguridad administrados puede tener consecuencias no deseadas. Es posible bloquear accidentalmente el tráfico necesario para que los clústeres funcionen correctamente y generar errores debido a que los nodos estén inaccesibles. Planifique y pruebe cuidadosamente las configuraciones de los grupos de seguridad antes de su implementación.

Solo puede especificar grupos de seguridad al crear un clúster. No se pueden añadir a un clúster ni a las instancias de clúster mientras se está ejecutando un clúster, pero es posible editar, añadir y eliminar reglas de los grupos de seguridad existentes. Las reglas surtirán efecto tan pronto como se guarden.

Los grupos de seguridad son restrictivos de forma predeterminada. A menos que se añada una regla que permita el tráfico, el tráfico se rechaza. Si existe más de una regla que se aplica al mismo tráfico y al mismo origen, se aplica la regla más permisiva. Por ejemplo, si tiene una regla que permite el tráfico SSH desde la dirección IP 192.0.2.12/32 y otra regla que permite el acceso a todo el tráfico TCP desde el rango 192.0.2.0/24, tiene prioridad la regla que permite todo el tráfico TCP desde el rango que incluye 192.0.2.12. En este caso, el cliente en la dirección 192.0.2.12 podría tener más acceso del deseado.

> **Important**
>
> Actúe con precaución al editar las reglas de grupos de seguridad para abrir puertos. Asegúrese de agregar reglas que solo permitan el tráfico desde los clientes de confianza y autenticados para los protocolos y puertos necesarios para ejecutar las cargas de trabajo.

Puede configurar Bloquear acceso público de Amazon EMR en cada región que utilice para evitar la creación de clústeres si una regla permite el acceso público en algún puerto que no haya agregado a una lista de excepciones. En el AWS caso de las cuentas creadas después de julio de 2019, Amazon EMR bloquea el acceso público y está activado de forma predeterminada. En el AWS caso de las cuentas que crearon un clúster antes de julio de 2019, el bloqueo de acceso público de Amazon EMR está desactivado de forma predeterminada. Para obtener más información, consulte Uso de Bloquear el acceso público de Amazon EMR.

Temas

- Uso de grupos de seguridad administrados por Amazon EMR
- Trabajo con grupos de seguridad adicionales para un clúster de Amazon EMR
- Especificación de los grupos de seguridad adicionales administrados por Amazon EMR
- Especificación EC2 de grupos de seguridad para EMR Notebooks
- Uso de Bloquear el acceso público de Amazon EMR

> **Note**
>
> Amazon EMR tiene como objetivo utilizar alternativas inclusivas para términos industriales potencialmente ofensivos o no inclusivos, como “maestro” y “esclavo”. Hemos adoptado una nueva terminología para fomentar una experiencia más inclusiva y que así pueda entender mejor los componentes del servicio.
> Ahora describimos los “nodos” como instancias y describimos los tipos de instancias de Amazon EMR como instancias principales, básicas y de tareas. Durante la transición, es posible que siga encontrando referencias antiguas a términos obsoletos, como los que se refieren a los grupos de seguridad de Amazon EMR.

Uso de grupos de seguridad administrados por Amazon EMR

> **Note**
>
> Amazon EMR tiene como objetivo utilizar alternativas inclusivas para términos industriales potencialmente ofensivos o no inclusivos, como “maestro” y “esclavo”. Hemos adoptado una nueva terminología para fomentar una experiencia más inclusiva y que así pueda entender mejor los componentes del servicio.
> Ahora describimos los “nodos” como instancias y describimos los tipos de instancias de Amazon EMR como instancias principales, básicas y de tareas. Durante la transición, es posible que siga encontrando referencias antiguas a términos obsoletos, como los que se refieren a los grupos de seguridad de Amazon EMR.

Son varios los grupos de seguridad administrados que están asociados a la instancia principal y con las instancias secundarias y de tareas de un clúster. Se requiere un grupo de seguridad administrado adicional para el acceso al servicio al crear un clúster en una subred privada. Para obtener más información sobre la función de los grupos de seguridad administrados con respecto a la configuración de la red, consulte Opciones de Amazon VPC al lanzar un clúster.

Cuando especifique grupos de seguridad administrados para un clúster, debe utilizar el mismo tipo de grupo de seguridad, predeterminado o personalizado, para todos los grupos. Por ejemplo, no puede especificar un grupo de seguridad personalizado para la instancia principal y, acto seguido, no especificar un grupo de seguridad personalizado para las instancias secundarias y de tareas.

Si piensa utilizar los grupos de seguridad administrados predeterminados, no es necesario que los especifique al crear un clúster. Amazon EMR utiliza automáticamente los valores predeterminados. Además, si los valores predeterminados aún no existen en la VPC del clúster, Amazon EMR los crea. Amazon EMR también los crea si los especifica de forma explícita y aún no existen.

Puede editar reglas en los grupos de seguridad administrados una vez que se hayan creado los clústeres. Cuando se crea un clúster nuevo, Amazon EMR comprueba las reglas de los grupos de seguridad administrados que se especifican y, a continuación, crea las reglas entrantes que faltan y que el clúster nuevo necesita, junto con las reglas que se hayan podido agregar anteriormente. A menos que se indique lo contrario, cada regla para los grupos de seguridad administrados por Amazon EMR predeterminados también se agrega a los grupos de seguridad administrados por Amazon EMR personalizados que especifique.

Los grupos de seguridad administrados predeterminados son los siguientes:

- ElasticMapReduce-principal

  Para ver las reglas de este grupo de seguridad, consulte Grupo de seguridad administrado por Amazon EMR para la instancia principal (subredes públicas).

- ElasticMapReduce-núcleo

  Para ver las reglas de este grupo de seguridad, consulte Grupo de seguridad administrado por Amazon EMR para las instancias básicas y de tareas (subredes públicas).

- ElasticMapReduce-Primario-Privado

  Para ver las reglas de este grupo de seguridad, consulte Grupo de seguridad administrado por Amazon EMR para la instancia principal (subredes privadas).

- ElasticMapReduce-Núcleo: privado

  Para ver las reglas de este grupo de seguridad, consulte Grupo de seguridad administrado por Amazon EMR para las instancias secundarias y de tareas (subredes privadas).

- ElasticMapReduce-ServiceAccess

  Para ver las reglas de este grupo de seguridad, consulte Grupo de seguridad administrado por Amazon EMR para el acceso de los servicios (subredes privadas).

Grupo de seguridad administrado por Amazon EMR para la instancia principal (subredes públicas)

El grupo de seguridad administrado predeterminado para la instancia principal en las subredes públicas tiene el nombre de ElasticMapReduce grupo -primary. Tiene las siguientes reglas: Si especifica un grupo de seguridad administrado personalizado, Amazon EMR agrega las mismas reglas a su grupo de seguridad personalizado.

Tipo	Protoc	Interva de puerto	Origen	Detalles
Reglas de entrada				

Tipo	Protoc	Interva de puertos	Origen	Detalles
Todos los ICMP-IPv4	Todos	N/A	El ID de grupo del grupo de seguridad administrado para la instancia principal. En otras palabras, el mismo grupo de seguridad en el que aparece la regla.	Estas reglas reflexivas permiten el tráfico entrante desde cualquier instancia asociada al grupo de seguridad especificado. El uso del grupo de seguridad predeterminado `ElasticMapReduce-primary` para varios clústeres permite que los nodos secundarios y de tareas de dichos clústeres se comuniquen entre sí a través de ICMP o de cualquier puerto TCP o UDP. Especifique grupos de seguridad administrados personalizados para restringir el acceso entre clústeres.
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		
Todos los ICMP-IPV4	Todos	N/A	El ID de grupo del grupo de seguridad administrado especificado para los nodos secundarios y de tareas.	Estas reglas permiten todo el tráfico ICMP entrante y el tráfico a través de cualquier puerto TCP o UDP desde cualquier instancia secundaria y de tareas asociada al grupo de seguridad especificado, incluso si las instancias se encuentran en clústeres distintos.
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		
Personalizada	TCP	8443	Diversos rangos de direcciones IP de Amazon	Estas reglas permiten que el administrador del clúster se comunique con el nodo principal.

Para conceder a los orígenes de confianza acceso SSH al grupo de seguridad principal con la consola

Para editar los grupos de seguridad, debe tener permiso para administrar los grupos de seguridad de la VPC en la que se encuentra el clúster. Para obtener más información, consulte Cambiar los

permisos de un usuario y el ejemplo de política que permite administrar grupos de EC2 seguridad en la Guía del usuario de IAM.

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Seleccione Clusters (Clústeres). Seleccione el ID del clúster que desea modificar.
3. En el panel Red y seguridad, expanda el menú desplegable de grupos EC2 de seguridad (firewall).
4. En Nodo principal, elija su grupo de seguridad.
5. Elija Editar reglas de entrada.
6. Compruebe si hay una regla de entrada que permita el acceso público con la siguiente configuración. Si existe, seleccione Eliminar para eliminarla.

   - Tipo

     SSH
   - Puerto

     22
   - Origen

     0.0.0.0/0 personalizado

   > **Warning**
   >
   > Antes de diciembre de 2020, había una regla preconfigurada para permitir el tráfico entrante en el puerto 22 desde todas las fuentes. Esta regla se creó para simplificar las conexiones SSH iniciales al nodo principal. Le recomendamos encarecidamente que elimine esta regla de entrada y que restrinja el tráfico a los orígenes de confianza.

7. Desplácese a la parte inferior de la lista de reglas y seleccione Agregar regla.
8. En Type (Tipo), seleccione SSH.

   Al seleccionar SSH, se ingresa automáticamente TCP en Protocolo y 22 en Rango de puertos.
9. Como origen, seleccione Mi IP para agregar automáticamente su dirección IP como dirección de origen. También puede agregar un rango de direcciones IP de clientes de confianza

personalizadas o crear reglas adicionales para otros clientes. Muchos entornos de red asignan direcciones IP de forma dinámica, por lo que es posible que en el futuro necesite actualizar las direcciones IP de los clientes de confianza.

10. Seleccione Save.
11. Si lo desea, puede seleccionar el otro grupo de seguridad en Nodos principales y de tareas en el panel Red y seguridad y repetir los pasos anteriores para permitir que el cliente SSH acceda a dichos nodos principales y de tareas.

Grupo de seguridad administrado por Amazon EMR para las instancias básicas y de tareas (subredes públicas)

El grupo de seguridad administrado predeterminado para las instancias principales y de tareas en las subredes públicas tiene el nombre de ElasticMapReduce grupo -core. El grupo de seguridad administrado predeterminado tiene las siguientes reglas, y Amazon EMR agrega las mismas reglas si especifica un grupo de seguridad administrado personalizado.

Tipo	Protoc	Interva de puerto	Origen	Detalles
Reglas de entrada				
Todos los ICMP-IPV4	Todos	N/A	El ID de grupo del grupo de seguridad administrado para las instancias secundarias y de tareas. En otras palabras, el mismo grupo de seguridad en el que aparece la regla.	Estas reglas reflexivas permiten el tráfico entrante desde cualquier instancia asociada al grupo de seguridad especificado. El uso del grupo de seguridad predeterm inado `ElasticMapReduce-core` para varios clústeres permite que las instancias secundarias y de tareas de dichos clústeres se comuniquen entre sí a través de ICMP o de cualquier puerto TCP o UDP. Especifique grupos de seguridad administrados personali zados para restringir el acceso entre clústeres .
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		

Tipo	Protoc	Interva de puertos	Origen	Detalles
Todos los ICMP-IPV4	Todos	N/A	El ID de grupo del grupo de seguridad administrado para la instancia principal.	Estas reglas permiten todo el tráfico ICMP entrante y el tráfico a través de cualquier puerto TCP o UDP desde cualquier instancia principal asociada al grupo de seguridad especificado, incluso si las instancias se encuentran en clústeres distintos.
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		

Grupo de seguridad administrado por Amazon EMR para la instancia principal (subredes privadas)

El grupo de seguridad administrado predeterminado para la instancia principal en las subredes privadas tiene el nombre de grupo -Primary-Private. ElasticMapReduce El grupo de seguridad administrado predeterminado tiene las siguientes reglas, y Amazon EMR agrega las mismas reglas si especifica un grupo de seguridad administrado personalizado.

Tipo	Protoc	Interva de puertos	Origen	Detalles
Reglas de entrada				
Todos los ICMP-IPv4	Todos	N/A	El ID de grupo del grupo de seguridad administrado para la instancia principal. En otras palabras, el mismo grupo de seguridad en el que aparece la regla.	Estas reglas reflexivas permiten el tráfico entrante desde cualquier instancia asociada al grupo de seguridad especificado y que esté accesible desde dentro de la subred privada. El uso del grupo de seguridad predeterm inado `ElasticMapReduce-Primary-Private` para varios clústeres permite que los nodos secundarios y de tareas de dichos clústeres se comuniquen entre sí a través
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		

Tipo	Protoc	Interva de puertos	Origen	Detalles
				de ICMP o de cualquier puerto TCP o UDP. Especifique grupos de seguridad administr ados personalizados para restringir el acceso entre clústeres.
Todos los ICMP-IPV4	Todos	N/A	El ID de grupo del grupo de seguridad administrado para los nodos secundarios y de tareas.	Estas reglas permiten todo el tráfico ICMP entrante y el tráfico a través de cualquier puerto TCP o UDP desde cualquier instancia secundaria y de tareas asociada al grupo de seguridad especificado y que esté accesible desde la subred privada, incluso si las instancias se encuentran en clústeres distintos.
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		
HTTPS (8443)	TCP	8443	El ID de grupo del grupo de seguridad administrado para el acceso de los servicios en una subred privada.	Esta regla permite que el administrador del clúster se comunique con el nodo principal.
Reglas de salida				
Todo el tráfico	Todos	Todos	0.0.0.0/0	Permite el acceso de salida a Internet.

Tipo	Protoc	Interva de puertos	Origen	Detalles
TCP personali zada	TCP	9443	El ID de grupo del grupo de seguridad administrado para el acceso de los servicios en una subred privada.	Si se elimina la regla de salida predeterm inada “Todo el tráfico” anterior, esta regla es un requisito mínimo para las versiones 5.30.0 y posteriores de Amazon EMR. **Note** Amazon EMR no agrega esta regla cuando utiliza un grupo de seguridad administrado personalizado.
TCP personali zada	TCP	80 (http) o 443 (https)	El ID de grupo del grupo de seguridad administrado para el acceso de los servicios en una subred privada.	Si se elimina la regla de salida predeterm inada “Todo el tráfico” anterior, esta regla es un requisito mínimo para que las versiones 5.30.0 y posteriores de Amazon EMR se conecten a Amazon S3 a través de https. **Note** Amazon EMR no agrega esta regla cuando utiliza un grupo de seguridad administrado personalizado.

Grupo de seguridad administrado por Amazon EMR para las instancias secundarias y de tareas (subredes privadas)

El grupo de seguridad administrado predeterminado para las instancias principales y de tareas en las subredes privadas tiene el nombre de grupo -Core-Private. ElasticMapReduce El grupo de seguridad administrado predeterminado tiene las siguientes reglas, y Amazon EMR agrega las mismas reglas si especifica un grupo de seguridad administrado personalizado.

Tipo	Protoc	Interva de puertos	Origen	Detalles
Reglas de entrada				
Todos los ICMP-IPV4	Todos	N/A	El ID de grupo del grupo de seguridad administrado para las instancias secundarias y de tareas. En otras palabras, el mismo grupo de seguridad en el que aparece la regla.	Estas reglas reflexivas permiten el tráfico entrante desde cualquier instancia asociada al grupo de seguridad especificado. El uso del grupo de seguridad predeterm inado `ElasticMapReduce-core` para varios clústeres permite que las instancias secundarias y de tareas de dichos clústeres se comuniquen entre sí a través de ICMP o de cualquier puerto TCP o UDP. Especifique grupos de seguridad administrados personali zados para restringir el acceso entre clústeres .
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		
Todos los ICMP-IPV4	Todos	N/A	El ID de grupo del grupo de seguridad administrado para la instancia principal.	Estas reglas permiten todo el tráfico ICMP entrante y el tráfico a través de cualquier puerto TCP o UDP desde cualquier instancia principal asociada al grupo de seguridad especificado, incluso si las instancias se encuentran en clústeres distintos.
Todos los TCP	TCP	Todos		
Todo el UDP	UDP	Todos		
HTTPS (8443)	TCP	8443	El ID de grupo del grupo de seguridad administrado para el acceso de los servicios en una subred privada.	Esta regla permite que el administrador del clúster se comunique con los nodos secundarios y de tareas.
Reglas de salida				

Tipo	Protoc	Interva de puertos	Origen	Detalles
Todo el tráfico	Todos	Todos	0.0.0.0/0	Consulte Edición de reglas de salida a continuación.
TCP personalizada	TCP	80 (http) o 443 (https)	El ID de grupo del grupo de seguridad administrado para el acceso de los servicios en una subred privada.	Si se elimina la regla de salida predeterminada “Todo el tráfico” anterior, esta regla es un requisito mínimo para que las versiones 5.30.0 y posteriores de Amazon EMR se conecten a Amazon S3 a través de https. **Note** Amazon EMR no agrega esta regla cuando utiliza un grupo de seguridad administrado personalizado.

Edición de reglas de salida

De forma predeterminada, Amazon EMR crea este grupo de seguridad con reglas de salida que permiten todo el tráfico saliente en todos los protocolos y puertos. Se selecciona Permitir todo el tráfico saliente porque varias aplicaciones cliente y de Amazon EMR que se pueden ejecutar en clústeres de Amazon EMR pueden requerir reglas de salida diferentes. Amazon EMR no puede anticipar estos ajustes específicos al crear grupos de seguridad predeterminados. Puede limitar las salidas en sus grupos de seguridad para incluir solo las reglas que se adapten a sus casos de uso y políticas de seguridad. Como mínimo, este grupo de seguridad requiere las siguientes reglas de salida, pero es posible que algunas aplicaciones necesiten una salida adicional.

Tipo	Protoc	Rango de puerto	Destino	Detalles
Todos los TCP	TCP	Todos	pl- *xxxxxxxx*	Lista de prefijos de Amazon S3 administrados `com.amazonaws.` *`MyRegion`*`.s3`.

Tipo	Protoc	Rango de puerto	Destino	Detalles
All Traffic	Todos	Todos	sg- *xxxxxxxxx xxxxxxxx*	El ID del grupo de seguridad `ElasticMapReduce-Core-Private` .
All Traffic	Todos	Todos	sg- *xxxxxxxxx xxxxxxxx*	El ID del grupo de seguridad `ElasticMapReduce-Primary-Private` .
TCP personalizada	TCP	9443	sg- *xxxxxxxxx xxxxxxxx*	El ID del grupo de seguridad `ElasticMapReduce-ServiceAccess` .

Grupo de seguridad administrado por Amazon EMR para el acceso de los servicios (subredes privadas)

El grupo de seguridad administrado predeterminado para el acceso a los servicios en subredes privadas tiene el nombre de grupo de ElasticMapReduce- ServiceAccess. Tiene reglas de entrada y reglas de salida que permiten el tráfico a través de HTTPS (puerto 8443 o 9443) hacia los demás grupos de seguridad administrados en las subredes privadas. Estas reglas permiten que el administrador del clúster se comunique con el nodo principal y con nodos principales y de tarea. Se necesitan las mismas reglas si utiliza grupos de seguridad personalizados.

Tipo	Protoc	Interva de puertos	Origen	Detalles
Reglas de entrada necesarias para clústeres de Amazon EMR con versiones 5.30.0 y posteriores.				
TCP personalizada	TCP	9443	El ID de grupo del grupo de seguridad administrado de la instancia principal.	Esta regla permite la comunicación entre el grupo de seguridad de la instancia principal y el grupo de seguridad de acceso al servicio.
Reglas de salida necesarias para todos los clústeres de Amazon EMR				

Tipo	Protoc	Interva de puertos	Origen	Detalles
TCP personalizada	TCP	8443	El ID de grupo del grupo de seguridad administrado de la instancia principal.	Estas reglas permiten que el administrador del clúster se comunique con el nodo principal y con nodos principales y de tarea.
TCP personalizada	TCP	8443	El ID de grupo del grupo de seguridad administrado para las instancias secundarias y de tareas.	Estas reglas permiten que el administrador del clúster se comunique con el nodo principal y con nodos principales y de tarea.

Trabajo con grupos de seguridad adicionales para un clúster de Amazon EMR

Tanto si utiliza los grupos de seguridad administrados predeterminados como si especifica grupos de seguridad administrados personalizados, puede utilizar grupos de seguridad adicionales. Los grupos de seguridad adicionales le proporcionan la flexibilidad necesaria para adaptar el acceso entre diferentes clústeres y desde clientes, aplicaciones y recursos externos.

Considere el siguiente escenario de ejemplo. Dispone de varios clústeres y necesita que se comuniquen entre ellos, pero desea permitir el acceso SSH entrante a la instancia principal solo a un determinado subconjunto de clústeres. Para ello, puede utilizar el mismo conjunto de grupos de seguridad administrados para los clústeres. A continuación, debe crear grupos de seguridad adicionales que permitan el acceso SSH entrante desde los clientes de confianza y especificar los grupos de seguridad adicionales para la instancia principal de cada uno de los clústeres del subconjunto.

Puede aplicar hasta 15 grupos de seguridad adicionales a la instancia principal, 15 a las instancias secundarias y de tareas, y 15 para el acceso al servicio (en subredes privadas). Si fuera necesario, puede especificar el mismo grupo de seguridad adicional para las instancias principales, las instancias secundarias y de tareas y el acceso al servicio. El número máximo de grupos de seguridad

y reglas de la cuenta está sujeto a los límites de la cuenta. Para obtener más información, consulte Límites de los grupos de seguridad en la Guía del usuario de Amazon VPC.

Especificación de los grupos de seguridad adicionales administrados por Amazon EMR

Puede especificar grupos de seguridad mediante la AWS Management Console AWS CLI, la o la API de Amazon EMR. Si no especifica grupos de seguridad, Amazon EMR crea grupos de seguridad predeterminados. La especificación de grupos de seguridad adicionales es opcional. Puede asignar grupos de seguridad adicionales a las instancias principales, a las instancias secundarias y de tareas y al acceso de los servicios (solo para las subredes privadas).

Console

Para especificar grupos de seguridad con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Redes, seleccione la flecha situada junto a los grupos EC2 de seguridad (firewall) para ampliar esta sección. En Nodo principal y en Nodos principales y de tareas, se seleccionan de forma predeterminada los grupos de seguridad administrados por Amazon EMR predeterminados. Si utiliza una subred privada, también tiene la opción de seleccionar un grupo de seguridad para Acceso de servicio.
4. Para cambiar el grupo de seguridad administrado por Amazon EMR, utilice el menú desplegable Elegir los grupos de seguridad para seleccionar una opción diferente de la lista de opciones del grupo de seguridad administrado por Amazon EMR. Dispone de un grupo de seguridad administrado por Amazon EMR tanto para Nodo principal como para Nodos principales y de tareas.
5. Para agregar grupos de seguridad personalizados, utilice el mismo menú desplegable Elegir los grupos de seguridad para seleccionar hasta cuatro grupos de seguridad personalizados de la lista de opciones Grupos de seguridad personalizados. Puede tener hasta cuatro grupos de seguridad personalizados tanto para Nodo principal como para Nodos principales y de tareas.
6. Elija cualquier otra opción que se aplique a su clúster.
7. Para lanzar el clúster, elija Crear clúster.

Especificación de grupos de seguridad con la AWS CLI

Para especificar grupos de seguridad mediante el, AWS CLI utilice el `create-cluster` comando con los siguientes parámetros de la `--ec2-attributes` opción:

Parámetro	Descripción
`EmrManagedPrimarySecurityGroup`	Utilice este parámetro para especificar un grupo de seguridad administrado personalizado para la instancia principal. Si se especifica este parámetro, también se deben especificar `EmrManagedCoreSecurityGroup`. Para clústeres en subredes privadas, también se debe especificar `ServiceAccessSecurityGroup`.
`EmrManagedCoreSecurityGroup`	Utilice este parámetro para especificar un grupo de seguridad administrado personalizado para las instancias secundarias y de tareas. Si se especifica este parámetro, también se deben especificar `EmrManagedPrimarySecurityGroup`. Para clústeres en subredes privadas, también se debe especificar `ServiceAccessSecurityGroup`.
`ServiceAccessSecurityGroup`	Utilice este parámetro para especificar un grupo de seguridad administrado personalizado para el acceso de los servicios, que se aplica únicamente a los clústeres de las subredes privadas. El grupo de seguridad que especifique como `ServiceAccessSecurityGroup` no debe usarse para ningún otro propósito y también debe reservarse para Amazon EMR. Si se especifica este parámetro, también se

Parámetro	Descripción
	deben especificar `EmrManagedPrimarySecurityGroup` .
`AdditionalPrimarySecurityGroups`	Utilice este parámetro para especificar hasta cuatro grupos de seguridad adicionales para la instancia principal.
`AdditionalCoreSecurityGroups`	Utilice este parámetro para especificar hasta cuatro grupos de seguridad adicionales para las instancias secundarias y de tareas.

Example : especifique grupos de seguridad administrados por Amazon EMR personalizados y grupos de seguridad adicionales

En el siguiente ejemplo, se especifican grupos de seguridad administrados por Amazon EMR personalizados para un clúster en una subred privada, varios grupos de seguridad adicionales para la instancia principal y un único grupo de seguridad adicional para las instancias secundarias y de tareas.

> **Note**
>
> Se incluyen caracteres de continuación de línea de Linux (\) para facilitar la lectura. Se pueden eliminar o utilizar en los comandos de Linux. En Windows, elimínelos o sustitúyalos por un signo de intercalación (^).

```
aws emr create-cluster --name "ClusterCustomManagedAndAdditionalSGs" \
--release-label emr-emr-7.10.0 --applications Name=Hue Name=Hive \
Name=Pig --use-default-roles --ec2-attributes \
SubnetIds=subnet-xxxxxxxxxxxx,KeyName=myKey,\
ServiceAccessSecurityGroup=sg-xxxxxxxxxxxx,\
EmrManagedPrimarySecurityGroup=sg-xxxxxxxxxxxx,\
EmrManagedCoreSecurityGroup=sg-xxxxxxxxxxx,\
AdditionalPrimarySecurityGroups=['sg-xxxxxxxxxxx',\
'sg-xxxxxxxxxxx','sg-xxxxxxxxxx'],\
AdditionalCoreSecurityGroups=sg-xxxxxxxxxxx \
```

```
--instance-type m5.xlarge
```

Para obtener más información, consulte create-cluster en la Referencia de comandos de la AWS CLI .

Especificación EC2 de grupos de seguridad para EMR Notebooks

Al crear un cuaderno de EMR, se utilizan dos grupos de seguridad para controlar el tráfico de red entre el cuaderno de EMR y el clúster de Amazon EMR cuando se usa el editor de cuadernos. Los grupos de seguridad predeterminados tienen reglas mínimas que únicamente permiten el tráfico de red entre el servicio Cuadernos de Amazon EMR y los clústeres a los que están asociados los cuadernos.

Un cuaderno EMR utiliza Apache Livy para comunicarse con el clúster mediante un proxy a través del puerto TCP 18888. Cuando crea grupos de seguridad personalizados con reglas adaptadas al entorno, puede limitar el tráfico de la red de forma que solo un subconjunto de cuadernos pueda ejecutar código en el editor de cuadernos en determinados clústeres. El clúster utiliza su seguridad personalizada además de los grupos de seguridad predeterminados del clúster. Para obtener más información, consulte Control del tráfico de red con grupos de seguridad en la Guía de administración de Amazon EMR y Especificación EC2 de grupos de seguridad para EMR Notebooks.

Grupo EC2 de seguridad predeterminado para la instancia principal

El grupo EC2 de seguridad predeterminado de la instancia principal está asociado a la instancia principal además de a los grupos de seguridad del clúster de la instancia principal.

Nombre del grupo: ElasticMapReduceEditors-Livy

Reglas

- Entrada

  Permita el puerto TCP 18888 desde cualquier recurso del grupo de EC2 seguridad predeterminado para EMR Notebooks

- Salida

  Ninguno

Grupo EC2 de seguridad predeterminado para EMR Notebooks

El grupo EC2 de seguridad predeterminado del bloc de notas de EMR está asociado al editor de bloc de notas de cualquier bloc de notas de EMR al que esté asignado.

Nombre del grupo: -Editor ElasticMapReduceEditors

Reglas

- Entrada

  Ninguno
- Salida

  Permita que el puerto TCP 18888 acceda a cualquier recurso del grupo de EC2 seguridad predeterminado para los Notebooks EMR.

Grupo EC2 de seguridad personalizado para los cuadernos EMR al asociar cuadernos con los repositorios de Git

Para vincular un repositorio de Git a su cuaderno, el grupo de seguridad del cuaderno de EMR debe incluir una regla de salida para permitir que el cuaderno envíe tráfico a Internet. Se recomienda crear un nuevo grupo de seguridad para este fin. Al actualizar el grupo de seguridad ElasticMapReduceEditors-Editor predeterminado, es posible que se apliquen las mismas reglas de salida a otros cuadernos adjuntos a este grupo de seguridad.

Reglas

- Entrada

  Ninguno
- Salida

  Permita que el cuaderno envíe tráfico a Internet a través del clúster, como se muestra en el siguiente ejemplo: El valor 0.0.0.0/0 se usa solo como ejemplo. Puede modificar esta regla para especificar las direcciones IP de sus repositorios basados en Git.

Tipo	Protocolo	Rango de puerto	Destino

Tipo	Protocolo	Rango de puerto	Destino
Regla TCP personalizada	TCP	18888	SG-
HTTPS	TCP	443	0.0.0.0/0

Uso de Bloquear el acceso público de Amazon EMR

Bloquear el acceso público (BPA) de Amazon EMR le impide lanzar un clúster en una subred pública si el clúster tiene una configuración de seguridad que permite el tráfico entrante desde direcciones IP públicas en un puerto.

> **Important**
>
> Bloquear el acceso público está habilitado de forma predeterminada. Si desea aumentar la protección de la cuenta, le recomendamos que lo mantenga habilitado.

Explicación de Bloquear el acceso público

Puede utilizar la configuración de cuenta de Bloquear el acceso público para administrar de forma centralizada el acceso a la red pública a los clústeres de Amazon EMR.

Cuando un usuario de su empresa Cuenta de AWS lanza un clúster, Amazon EMR comprueba las reglas de puerto del grupo de seguridad del clúster y las compara con las reglas de tráfico entrante. Si el grupo de seguridad tiene una regla de entrada que abre los puertos a las direcciones IP públicas IPv4 0.0.0.0/0 o IPv6 : :/0, y esos puertos no se especifican como excepciones para su cuenta, Amazon EMR no permite que el usuario cree el clúster.

Si un usuario modifica las reglas del grupo de seguridad de un clúster en ejecución en una subred pública para tener una regla de acceso público que infrinja la configuración de BPA de su cuenta, Amazon EMR revoca la nueva regla si tiene permiso para hacerlo. Si Amazon EMR no tiene permiso para revocar la regla, crea un evento en el panel de AWS Health que describe la infracción. Para conceder el permiso de revocación de la regla a Amazon EMR, consulte Configuración de Amazon EMR para revocar las reglas de los grupos de seguridad.

Bloquear acceso público está habilitado de forma predeterminada para todos los clústeres de cada Región de AWS de su Cuenta de AWS. BPA se aplica a todo el ciclo de vida de un clúster, pero no a los clústeres que se crean en subredes privadas. Puede configurar excepciones a la regla de BPA; el puerto 22 es una excepción de forma predeterminada. Para obtener más información sobre la configuración de excepciones, consulte Configuración de Bloquear el acceso público.

Configuración de Bloquear el acceso público

Puede actualizar los grupos de seguridad y la configuración de Bloquear el acceso público en sus cuentas en cualquier momento.

Puede activar y desactivar la configuración de bloqueo de acceso público (BPA) con AWS Management Console, AWS Command Line Interface (AWS CLI) y la API Amazon EMR. La configuración se aplica a toda su cuenta de forma puntual. Region-by-Region Para mantener la seguridad del clúster, le recomendamos usar BPA.

Console

Para configurar: bloqueo de acceso público con la consola

1. Inicie sesión en y, a continuación AWS Management Console, abra la consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En la barra de navegación superior, seleccione la Región que quiera configurar si aún no está seleccionada.
3. En EMR activado, EC2 en el panel de navegación izquierdo, selecciona Bloquear acceso público.
4. En Block public access settings (Configuración de Block Public Access), complete los pasos siguientes.

Para…	Haga lo siguiente...
Activar o desactivar Block Public Access	Seleccione Editar, Activar o Desactiva r según corresponda y, a continuación, seleccione Guardar.
Editar puertos en la lista de excepciones	1.

Para…	Haga lo siguiente...
	Seleccione Editar y busque la sección Excepciones de rango de puertos. 2. Para añadir puertos a la lista de excepciones, elija Add a port range (Añadir un rango de puertos) y escriba un nuevo puerto o rango de puertos. Repita para cada puerto o rango de puertos que desee añadir. 3. Para eliminar un puerto o rango de puertos, elija Eliminar, junto a la entrada en la lista de rangos de puertos. 4. Seleccione Save.

AWS CLI

Para configurar el bloqueo del acceso público mediante el AWS CLI

- Utilice el comando `aws emr put-block-public-access-configuration` para configurar Block Public Access, tal y como se muestra en los siguientes ejemplos.

Para…	Haga lo siguiente...
Activar Block Public Access	Defina `BlockPublicSecurityGroupRules` en `true` como se muestra en el ejemplo siguiente. Para que el clúster se lance, ningún grupo de seguridad asociado a un clúster puede tener una regla de entrada que permita el acceso público. `aws emr put-block-public-access-configuration --block-public-access-configuration BlockPublicSecurityGroupRules=true`
Desactivar Block Public Access	Defina `BlockPublicSecurityGroupRules` en `false` como se muestra en el ejemplo siguiente. Los grupos de seguridad asociados a un clúster pueden tener reglas de entrada que permitan el acceso público en cualquier puerto. No recomendamos esta configuración. `aws emr put-block-public-access-configuration --block-public-access-configuration BlockPublicSecurityGroupRules=false`

Para…	Haga lo siguiente...
Activar Block Public Access y especificar los puertos como excepciones	En el siguiente ejemplo se activa el bloqueo del acceso público y se especifica el puerto 22 y los puertos 100-101 como excepcion es. Esto permite crear clústeres si un grupo de seguridad asociado tiene una regla de entrada que permite el acceso público en los puertos 22, 100 o 101. `aws emr put-block-public-access-configuration --block-public-acc ess-configuration '{ "BlockPub licSecurityGroupRules": true, "PermittedPublicSecurityGro upRuleRanges": [ { "MinRange": 22, "MaxRange": 22 }, { "MinRange ": 100, "MaxRange": 101 } ] }'`

Configuración de Amazon EMR para revocar las reglas de los grupos de seguridad

Amazon EMR necesita permiso para revocar las reglas de los grupos de seguridad y cumplir con la configuración de Bloquear el acceso público. Puede seguir uno de estos métodos para conceder a Amazon EMR el permiso que necesita:

- Recomendado Asocie la política administrada `AmazonEMRServicePolicy_v2` al rol de servicio. Para obtener más información, consulte Rol de servicio para Amazon EMR (rol de EMR).
- Cree una nueva política insertada que permita realizar la acción `ec2:RevokeSecurityGroupIngress` en los grupos de seguridad. Para obtener más información sobre cómo modificar una política de permisos de roles, consulte Modificación de una política de permisos de rol con la consola de IAM, la API de AWS y AWS CLI en la Guía del usuario de IAM.

Resolución de infracciones de Bloquear el acceso público

Si se produce una infracción de Bloquear el acceso público, puede mitigarla con una de las siguientes acciones:

- Si desea acceder a una interfaz web de su clúster, utilice una de las opciones descritas en Ver las interfaces web alojadas en clústeres de Amazon EMR para acceder a la interfaz a través de SSH (puerto 22).
- Para permitir el tráfico al clúster desde direcciones IP específicas en lugar de desde la dirección IP pública, agregue una regla de grupo de seguridad. Para obtener más información, consulte Añadir reglas a un grupo de seguridad en la Guía de EC2 introducción de Amazon.
- (No recomendado) Puede configurar las excepciones de BPA de Amazon EMR para que incluyan el puerto o el rango de puertos que desee. Al especificar una excepción de BPA, se crea un riesgo con un puerto desprotegido. Si planea especificar una excepción, debe eliminarla tan pronto como deje de ser necesaria. Para obtener más información, consulte Configuración de Bloquear el acceso público.

Identificación de los clústeres asociados a las reglas de los grupos de seguridad

Es posible que tenga que identificar todos los clústeres asociados a una regla de grupo de seguridad determinada o buscar la regla de grupo de seguridad de un clúster determinado.

- Si conoce el grupo de seguridad, puede identificar sus clústeres asociados si encuentra las interfaces de red del grupo de seguridad. Para obtener más información, consulta ¿Cómo puedo encontrar los recursos asociados a un grupo de EC2 seguridad de Amazon? no AWS re:Post. Las EC2 instancias de Amazon que estén conectadas a estas interfaces de red se etiquetarán con el ID del clúster al que pertenecen.
- Si desea buscar los grupos de seguridad de un clúster conocido, siga los pasos que se indican en Visualización del estado y los detalles del clúster de Amazon EMR. Puede encontrar los grupos de seguridad del clúster en el panel Red y seguridad de la consola o en el campo `Ec2InstanceAttributes` de la AWS CLI.

Validación de la conformidad de Amazon EMR

Los auditores externos evalúan la seguridad y el cumplimiento de Amazon EMR como parte de varios programas de AWS cumplimiento. Estos incluyen SOC, PCI, FedRAMP, HIPAA y otros.

Para obtener una lista de AWS los servicios incluidos en el ámbito de los programas de conformidad específicos, consulte los AWS servicios incluidos en el ámbito de aplicación por programa de conformidad. Para obtener información general, consulte Programas de conformidad de AWS.

Puede descargar informes de auditoría de terceros utilizando AWS Artifact. Para obtener más información, consulte Descargar informes en AWS Artifact.

La responsabilidad de conformidad al utilizar Amazon EMR se determina en función de la confidencialidad de los datos, los objetivos de conformidad de la empresa y la legislación y normativa aplicables. Si el uso de Amazon EMR está sujeto al cumplimiento de normas como HIPAA, PCI o FedRAMP, proporciona recursos que le ayudarán a: AWS

- Guías de inicio rápido sobre seguridad y conformidad: estas guías de implementación analizan las consideraciones arquitectónicas y proporcionan los pasos para implementar entornos básicos centrados en la seguridad y el cumplimiento. AWS
- Documento técnico sobre cómo diseñar una arquitectura basada en la seguridad y el cumplimiento de la HIPAA: este documento técnico describe cómo pueden utilizar las empresas para crear aplicaciones que cumplan con la HIPAA. AWS
- AWS Recursos de cumplimiento: esta colección de libros de trabajo y guías puede aplicarse a su sector y ubicación.
- AWS Config— Este AWS servicio evalúa en qué medida las configuraciones de sus recursos cumplen con las prácticas internas, las directrices del sector y las normativas.
- AWS Security Hub— Este AWS servicio proporciona una visión integral del estado de su seguridad AWS que le ayuda a comprobar el cumplimiento de los estándares y las mejores prácticas del sector de la seguridad.

Resiliencia en Amazon EMR

La infraestructura AWS global se basa en AWS regiones y zonas de disponibilidad. AWS Las regiones proporcionan varias zonas de disponibilidad aisladas y separadas físicamente, que están conectadas mediante redes de baja latencia, alto rendimiento y alta redundancia. Con las zonas de disponibilidad, puede diseñar y utilizar aplicaciones y bases de datos que realizan una conmutación por error automática entre zonas de disponibilidad sin interrupciones. Las zonas de disponibilidad tienen una mayor disponibilidad, tolerancia a errores y escalabilidad que las infraestructuras tradicionales de centros de datos únicos o múltiples.

Para obtener más información sobre AWS las regiones y las zonas de disponibilidad, consulte la infraestructura global.AWS

Además de la infraestructura AWS global, Amazon EMR ofrece varias funciones que ayudan a respaldar sus necesidades de respaldo y resiliencia de datos.

- Integración con Amazon S3 a través de EMRFS
- Soporte para varios nodos principales

Seguridad de la infraestructura de Amazon EMR

Como servicio gestionado, Amazon EMR está protegido por la seguridad de la red AWS global. Para obtener información sobre los servicios AWS de seguridad y cómo se AWS protege la infraestructura, consulte Seguridad AWS en la nube. Para diseñar su AWS entorno utilizando las mejores prácticas de seguridad de la infraestructura, consulte Protección de infraestructuras en un marco de buena AWS arquitectura basado en el pilar de la seguridad.

Utiliza las llamadas a la API AWS publicadas para acceder a Amazon EMR a través de la red. Los clientes deben admitir lo siguiente:

- Seguridad de la capa de transporte (TLS). Exigimos TLS 1.2 y recomendamos TLS 1.3.
- Conjuntos de cifrado con confidencialidad directa total (PFS) como DHE (Ephemeral Diffie-Hellman) o ECDHE (Elliptic Curve Ephemeral Diffie-Hellman). La mayoría de los sistemas modernos como Java 7 y posteriores son compatibles con estos modos.

Además, las solicitudes deben estar firmadas mediante un ID de clave de acceso y una clave de acceso secreta que esté asociada a una entidad principal de IAM. También puedes utilizar AWS Security Token Service (AWS STS) para generar credenciales de seguridad temporales para firmar solicitudes.

Temas

- Conexión a Amazon EMR mediante un punto de conexión de VPC de tipo interfaz

Conexión a Amazon EMR mediante un punto de conexión de VPC de tipo interfaz

Puede conectarse directamente a Amazon EMR mediante un punto de enlace de VPC de interfaz (AWS PrivateLink) en su Virtual Private Cloud (VPC) en lugar de conectarse a través de Internet. Cuando utiliza un punto final de VPC de interfaz, la comunicación entre su VPC y Amazon EMR se lleva a cabo íntegramente dentro de la red. AWS Cada punto final de la VPC está representado por una o más interfaces de red elásticas (ENIs) con direcciones IP privadas en las subredes de la VPC.

El punto de conexión de la VPC de la interfaz conecta la VPC directamente a Amazon EMR sin necesidad de una pasarela de Internet, un dispositivo NAT, una conexión de VPN o una conexión. AWS Direct Connect Las instancias de la VPC no necesitan direcciones IP públicas para comunicarse con la API de Amazon EMR.

Para utilizar Amazon EMR a través de la VPC, debe conectarse desde una instancia que esté dentro de la VPC o conectar su red privada a la VPC a través de una red privada virtual (VPN) de Amazon o AWS Direct Connect. Para obtener más información sobre Amazon VPN, consulte Conexiones VPN en la Guía del usuario de Amazon Virtual Private Cloud. Para obtener más información AWS Direct Connect, consulte Crear una conexión en la Guía del AWS Direct Connect usuario.

Puede crear un punto de enlace de VPC de interfaz para conectarse a Amazon EMR mediante la AWS consola o AWS Command Line Interface los comandos ().AWS CLI Para obtener más información, consulte Creación de un punto de conexión de interfaz.

Después de crear un punto de conexión de VPC de tipo interfaz, si habilita nombres de host DNS privados para el punto de conexión, el punto de conexión predeterminado de Amazon EMR se resuelve en el punto de conexión de VPC. El punto de conexión del nombre de servicio predeterminado de Amazon EMR tiene el siguiente formato:

```
elasticmapreduce.Region.amazonaws.com
```

Si no habilita nombres de host de DNS privados, Amazon VPC proporciona un nombre de punto de conexión de DNS que puede utilizar en el siguiente formato.

```
VPC_Endpoint_ID.elasticmapreduce.Region.vpce.amazonaws.com
```

Para obtener más información, consulte Puntos de conexión de VPC de interfaz (AWS PrivateLink) en la Guía del usuario de Amazon VPC.

Amazon EMR permite llamar a todas las acciones de la API en su VPC.

Puede adjuntar políticas de punto de conexión de VPC a un punto de conexión de VPC para controlar el acceso de las entidades principales de IAM. También puede asociar grupos de seguridad con un punto de conexión de VPC para controlar el acceso de entrada y salida en función del origen y el destino del tráfico de red, como un rango de direcciones IP. Para obtener más información, consulte Control del acceso a los servicios con puntos de conexión de VPC.

Creación de una política de punto de conexión de VPC para Amazon EMR

Puede crear una política para los puntos de conexión de VPC de Amazon para Amazon EMR y especificar lo siguiente:

- La entidad principal que puede o no puede realizar acciones
- Las acciones que se pueden realizar
- Los recursos en los que se pueden llevar a cabo las acciones

Para más información, consulte Control del acceso a los servicios con puntos de conexión de VPC en la Guía del usuario de Amazon VPC.

Example — Política de punto final de VPC para denegar todo acceso desde una cuenta específica AWS

La siguiente política de punto final de VPC deniega a la AWS cuenta *123456789012* todo acceso a los recursos que utilizan el punto final.

```
{
    "Statement": [
        {
            "Action": "*",
            "Effect": "Allow",
            "Resource": "*",
            "Principal": "*"
        },
        {
            "Action": "*",
            "Effect": "Deny",
            "Resource": "*",
            "Principal": {
                "AWS": [
```

```
                    "123456789012"
                ]
            }
        }
    ]
}
```

Example : política de punto de conexión de VPC para permitir el acceso de VPC solo a una entidad principal de IAM especificada (usuario).

La siguiente política de puntos finales de VPC permite el acceso total solo a un usuario *lijuan* de la cuenta. AWS *123456789012* A las demás entidades principales de IAM se les deniega el acceso a través del punto de enlace.

```
{
    "Statement": [
        {
            "Action": "*",
            "Effect": "Allow",
            "Resource": "*",
            "Principal": {
                "AWS": [
                    "arn:aws:iam::123456789012:user/lijuan"
                ]
            }
        }]
}
```

Example : política de punto de conexión de VPC para permitir operaciones de EMR de solo lectura.

La siguiente política de puntos finales de VPC permite que solo la AWS cuenta realice *123456789012* las acciones especificadas de Amazon EMR.

Las acciones especificadas proporcionan el equivalente al acceso de solo lectura para Amazon EMR. Las demás acciones en la VPC se deniegan para la cuenta especificada. A las demás cuentas se les deniega el acceso. Para obtener una lista de acciones de Amazon EMR, consulte Acciones, recursos y claves de condición de Amazon EMR.

```
{
    "Statement": [
        {
```

```
            "Action": [
                "elasticmapreduce:DescribeSecurityConfiguration",
                "elasticmapreduce:GetBlockPublicAccessConfiguration",
                "elasticmapreduce:ListBootstrapActions",
                "elasticmapreduce:ViewEventsFromAllClustersInConsole",
                "elasticmapreduce:ListSteps",
                "elasticmapreduce:ListInstanceFleets",
                "elasticmapreduce:DescribeCluster",
                "elasticmapreduce:ListInstanceGroups",
                "elasticmapreduce:DescribeStep",
                "elasticmapreduce:ListInstances",
                "elasticmapreduce:ListSecurityConfigurations",
                "elasticmapreduce:DescribeEditor",
                "elasticmapreduce:ListClusters",
                "elasticmapreduce:ListEditors"
            ],
            "Effect": "Allow",
            "Resource": "*",
            "Principal": {
                "AWS": [
                    "123456789012"
                ]
            }
        }
    ]
}
```

Example : política de punto de conexión de VPC que deniega el acceso a un clúster específico.

La siguiente política de puntos finales de VPC permite el acceso total a todas las cuentas y entidades principales, pero deniega el acceso de la AWS cuenta *123456789012* a las acciones realizadas en el clúster de Amazon EMR con el ID de clúster. *j-A1B2CD34EF5G* Se siguen permitiendo otras acciones de Amazon EMR que no admiten permisos de recursos para los clústeres. Para obtener una lista de acciones de Amazon EMR y su tipo de recurso correspondiente, consulte Acciones, recursos y claves de condición para Amazon EMR.

```
{
    "Statement": [
        {
            "Action": "*",
            "Effect": "Allow",
            "Resource": "*",
            "Principal": "*"
```

```
        },
        {
            "Action": "*",
            "Effect": "Deny",
            "Resource": "arn:aws:elasticmapreduce:us-west-2:123456789012:cluster/j-
A1B2CD34EF5G",
            "Principal": {
                "AWS": [
                    "123456789012"
                ]
            }
        }
    ]
}
```

Administración de clústeres de Amazon EMR

Después de haber lanzado el clúster, puede conectarlo y administrarlo. Amazon EMR proporciona un conjunto de herramientas que puede utilizar para ello. En esta sección, se proporcionan instrucciones para conectarse a su clúster, explicarle las tareas pendientes, monitorizar el trabajo mientras se realiza y solucionar problemas.

Temas

- Conectar a un clúster de Amazon EMR
- Envío del trabajo a un clúster de Amazon EMR
- Visualización y monitorización de un clúster de Amazon EMR a medida que realiza su trabajo
- Utilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes
- Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera.
- Clonación de un clúster de Amazon EMR utilizando la consola
- Automatización de los clústeres de Amazon EMR recurrentes con AWS Data Pipeline

Conectar a un clúster de Amazon EMR

Cuando ejecute un clúster de Amazon EMR, a menudo lo único que tendrá que hacer es ejecutar una aplicación para analizar los datos y, a continuación, recopilar la salida desde un bucket de Amazon S3. En otras ocasiones, es posible que desee interactuar con el nodo principal mientras que el clúster se está ejecutando. Por ejemplo, es posible que desee conectar con el nodo principal para ejecutar consultas interactivas, comprobar archivos de registro, depurar un problema con el clúster, monitorizar el rendimiento mediante una aplicación como Ganglia que se ejecuta en el nodo principal, etc. Las secciones siguientes describen técnicas que puede utilizar para conectarse al nodo principal.

En un clúster de EMR, el nodo principal es una EC2 instancia de Amazon que coordina las EC2 instancias que se ejecutan como nodos principales y de tareas. El nodo principal expone un nombre de DNS pública que puede utilizar para conectarse al mismo. De forma predeterminada, Amazon EMR crea reglas de grupo de seguridad para el nodo principal, los nodos secundarios y los nodos de tarea que determinan cómo se obtiene acceso a los nodos.

> **Note**
>
> Puede conectarse al nodo principal solo mientras se está ejecutando el clúster. Cuando el clúster termina, la EC2 instancia que actúa como nodo principal finaliza y deja de estar disponible. Para conectarse al nodo principal, también debe autenticarse en el clúster. Puedes usar Kerberos para la autenticación o especificar una EC2 clave privada de un par de claves de Amazon al lanzar el clúster. Para obtener más información sobre la configuración de Kerberos y, a continuación, conectar, consulte Uso de Kerberos para la autenticación con Amazon EMR. Al lanzar un clúster desde la consola, la clave privada del EC2 key pair de Amazon se especifica en la sección Seguridad y acceso de la página Crear clúster.

De forma predeterminada, el grupo de seguridad ElasticMapReduce -master no permite el acceso SSH entrante. Es posible que tenga que añadir una regla entrante que permita el acceso de SSH (TCP puerto 22) desde los orígenes a los que desea tener acceso. Para obtener más información sobre la modificación de las reglas de los grupos de seguridad, consulte Añadir reglas a un grupo de seguridad en la Guía del EC2 usuario de Amazon.

> **Important**
>
> No modifique las reglas restantes del grupo de seguridad ElasticMapReduce -master. La modificación de estas reglas podría interferir con la operación del clúster.

Temas

- Antes de conectarse a Amazon EMR: autorice el tráfico entrante
- Conexión al nodo principal del clúster de Amazon EMR mediante SSH

Antes de conectarse a Amazon EMR: autorice el tráfico entrante

Antes de conectarse a un clúster de Amazon EMR, debe autorizar el tráfico SSH entrante (puerto 22) procedente de clientes de confianza, como la dirección IP de su ordenador. Para ello, edite las reglas del grupo de seguridad administrado de los nodos a los que desee conectarse. Por ejemplo, las siguientes instrucciones muestran cómo añadir una regla de entrada para el acceso SSH al grupo de seguridad ElasticMapReduce -master predeterminado.

Para obtener más información acerca del uso de los grupos de seguridad con Amazon EMR, consulte Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

Console

Para conceder a las fuentes de confianza acceso SSH al grupo de seguridad principal con la consola

Para editar los grupos de seguridad, debe tener permiso para administrar los grupos de seguridad de la VPC en la que se encuentra el clúster. Para obtener más información, consulte Cambiar los permisos de un usuario y el ejemplo de política que permite administrar grupos de EC2 seguridad en la Guía del usuario de IAM.

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija el clúster que desee actualizar. Se abrirá la página de detalles del clúster. Se preseleccionará la pestaña Propiedades de esta página.
3. En Redes, en la pestaña Propiedades, seleccione la flecha situada junto a los grupos de EC2 seguridad (firewall) para ampliar esta sección. En Nodo principal, seleccione el enlace del grupo de seguridad. Esto abre la EC2 consola.
4. Elija la pestaña Reglas de entrada y, a continuación, Editar reglas de entrada.
5. Compruebe si hay una regla de entrada que permita el acceso público con la siguiente configuración. Si existe, seleccione Eliminar para eliminarla.
   - Tipo

     SSH
   - Puerto

     22
   - Origen

     0.0.0.0/0 personalizado

> **Warning**
>
> Antes de diciembre de 2020, el grupo de seguridad ElasticMapReduce -master tenía una regla preconfigurada para permitir el tráfico entrante en el puerto 22 desde todas las fuentes. Esta regla se creó para simplificar las conexiones SSH iniciales al nodo principal. Le recomendamos encarecidamente que elimine esta regla de entrada y que restrinja el tráfico a los orígenes de confianza.

6. Desplácese a la parte inferior de la lista de reglas y seleccione Agregar regla.
7. En Type (Tipo), seleccione SSH. Esta selección introduce automáticamente TCP en Protocolo y 22 en Rango de puertos.
8. Como origen, seleccione Mi IP para agregar automáticamente su dirección IP como dirección de origen. También puede agregar un rango de direcciones IP de clientes de confianza personalizadas o crear reglas adicionales para otros clientes. Muchos entornos de red asignan direcciones IP de forma dinámica, por lo que es posible que en el futuro necesite actualizar las direcciones IP de los clientes de confianza.
9. Seleccione Save.
10. Si lo desea, vuelva al paso 3, seleccione Nodos principales y de tarea y repita los pasos 4 a 8. Esto otorga acceso al cliente SSH a los nodos principales y de tarea.

Conexión al nodo principal del clúster de Amazon EMR mediante SSH

Secure Shell (SSH) es un protocolo de red que puede utilizar para crear una conexión segura a un equipo remoto. Después de realizar una conexión, el terminal en su equipo local se comporta como si se estuviera ejecutando en el equipo remoto. Los comandos que emita a nivel local se ejecutan en el equipo remoto y la salida de comandos desde el equipo remoto aparece en la ventana del terminal.

Cuando usas SSH con AWS, te conectas a una EC2 instancia, que es un servidor virtual que se ejecuta en la nube. Cuando se trabaja con Amazon EMR, el uso más común de SSH es conectarse a la EC2 instancia que actúa como nodo principal del clúster.

El uso de SSH para conectarse al nodo principal le ofrece la posibilidad de monitorizar e interactuar con el clúster. Puede emitir comandos de Linux en el nodo principal, ejecutar aplicaciones como, por ejemplo, Hive y Pig de forma interactiva, examinar directorios, leer archivos de registro, etc.

También puede crear un túnel en la conexión SSH para ver las interfaces web alojadas en el nodo principal. Para obtener más información, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

Para conectar con el nodo principal mediante SSH, necesita el nombre de DNS público del nodo principal. Además, el grupo de seguridad asociado al nodo principal debe tener una regla de entrada que permita el tráfico SSH (puerto TCP 22) desde un origen que incluya el cliente donde se origina la conexión SSH. Es posible que tenga que añadir una regla para permitir una conexión SSH de su cliente. Para obtener más información sobre la modificación de las reglas de los grupos de seguridad, consulte Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR Añadir reglas a un grupo de seguridad en la Guía del EC2 usuario de Amazon.

Recuperar el nombre de DNS público del nodo principal

Puede recuperar el nombre de DNS público principal utilizando la consola de Amazon EMR y la AWS CLI.

Console

Para recuperar el nombre de DNS público del nodo principal con la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, seleccione el clúster en el que desea recuperar el nombre de DNS público.
3. Anote el valor del DNS público del nodo principal en la sección Resumen de la página de detalles del clúster.

CLI

Para recuperar el nombre DNS público del nodo principal con la AWS CLI

1. Para recuperar el identificador del clúster, escriba el siguiente comando.

   ```
   aws emr list-clusters
   ```

   El resultado muestra los clústeres, incluido el clúster IDs. Tenga en cuenta el ID del clúster al que se está conectando.

```
"Status": {
    "Timeline": {
        "ReadyDateTime": 1408040782.374,
        "CreationDateTime": 1408040501.213
    },
    "State": "WAITING",
    "StateChangeReason": {
        "Message": "Waiting after step completed"
    }
},
"NormalizedInstanceHours": 4,
"Id": "j-2AL4XXXXXX5T9",
"Name": "My cluster"
```

2. Para obtener una lista de las instancias de clúster incluido el nombre de DNS público para el clúster, escriba uno de los siguientes comandos. *j-2AL4XXXXXX5T9*Sustitúyalo por el ID de clúster devuelto por el comando anterior.

```
aws emr list-instances --cluster-id j-2AL4XXXXXX5T9
```

O bien:

```
aws emr describe-cluster --cluster-id j-2AL4XXXXXX5T9
```

La salida enumera las instancias de clúster, incluidos nombres de DNS y las direcciones IP. Anote el valor para `PublicDnsName`.

```
"Status": {
    "Timeline": {
        "ReadyDateTime": 1408040779.263,
        "CreationDateTime": 1408040515.535
    },
    "State": "RUNNING",
    "StateChangeReason": {}
},
"Ec2InstanceId": "i-e89b45e7",
"PublicDnsName": "ec2-###-##-##-###.us-west-2.compute.amazonaws.com"

"PrivateDnsName": "ip-###-##-##-###.us-west-2.compute.internal",
"PublicIpAddress": "##.###.###.##",
```

```
"Id": "ci-12XXXXXXXXFMH",
"PrivateIpAddress": "###.##.#.###"
```

Para más información, consulte Comandos de Amazon EMR en la AWS CLI.

Conéctate al nodo principal mediante SSH y una clave EC2 privada de Amazon en Linux, Unix y Mac OS X

Para crear una conexión SSH autenticada con un archivo de clave privada, debes especificar la clave privada del par de EC2 claves de Amazon al lanzar un clúster. Para obtener más información sobre cómo acceder a tu par de claves, consulta los pares de EC2 claves de Amazon en la Guía del EC2 usuario de Amazon.

Su equipo Linux muy probablemente incluye un cliente SSH de forma predeterminada. Por ejemplo, OpenSSH está instalado en la mayoría de los sistemas operativos Linux, Unix y macOS. Puede comprobar si tiene un cliente SSH escribiendo ssh en la línea de comandos. Si su equipo no reconoce el comando, instale un cliente SSH para conectarse al nodo principal. El proyecto OpenSSH ofrece una implementación gratuita de toda la suite de herramientas de SSH. Para obtener más información, consulte el sitio web OpenSSH.

Las siguientes instrucciones muestran la apertura de una conexión SSH al nodo principal de Amazon EMR en Linux, Unix y Mac OS X.

Para configurar los permisos de archivo de clave privada del par de claves

Antes de poder usar la EC2 clave privada del par de claves de Amazon para crear una conexión SSH, debes configurar los permisos en el `.pem` archivo para que solo el propietario de la clave tenga permiso para acceder al archivo. Esto es necesario para crear una conexión SSH mediante el terminal o el. AWS CLI

1. Asegúrese de permitir el tráfico SSH entrante. Para obtener instrucciones, consulte Antes de conectarse a Amazon EMR: autorice el tráfico entrante.
2. Localice su archivo `.pem`. Estas instrucciones suponen que el archivo se denomina `mykeypair.pem` y que se almacena en el directorio de inicio del usuario actual.
3. Escriba el comando siguiente para definir los permisos. *`~/mykeypair.pem`*Sustitúyalo por la ruta completa y el nombre del archivo de clave privada del par de claves. Por ejemplo, `C:/Users/<username>/.ssh/mykeypair.pem`.

```
chmod 400 ~/mykeypair.pem
```

Si no establece permisos en el archivo `.pem`, recibirá un error indicando que el archivo de clave no está protegido y la clave se rechazará. Para conectarse, solo tiene que establecer permisos en el archivo de clave privada del par de claves la primera vez que lo use.

Para conectarse al nodo principal utilizando el terminal

1. Abra una ventana de terminal. En Mac OS X, elija Aplicaciones > Utilidades > Terminal. En otras distribuciones de Linux, el terminal normalmente se encuentra en Aplicaciones > Accesorios > Terminal.
2. Para establecer una conexión al nodo principal, escriba el siguiente comando. *`ec2-###-##-##-###.compute-1.amazonaws.com`*Sustitúyalo por el nombre DNS público principal del clúster y *`~/mykeypair.pem`* sustitúyelo por la ruta completa y el nombre del `.pem` archivo. Por ejemplo, `C:/Users/<username>/.ssh/mykeypair.pem`.

```
ssh hadoop@ec2-###-##-##-###.compute-1.amazonaws.com -i ~/mykeypair.pem
```

> **Important**
>
> Debe utilizar el nombre de inicio de sesión `hadoop` cuando se conecte al nodo principal de Amazon EMR; de lo contrario, es posible que vea un error similar a `Server refused our key`.

3. Una advertencia indica que la autenticidad del host al que se está conectando no se puede verificar. Escriba `yes` para continuar.
4. Cuando haya terminado de trabajar en el nodo principal, escriba el siguiente comando para cerrar la conexión SSH.

```
exit
```

Si tiene dificultades para usar SSH para conectarse a su nodo principal, consulte Solucionar problemas de conexión a su instancia.

Conectarse al nodo principal mediante SSH en Windows

Los usuarios de Windows pueden utilizar un cliente SSH como PuTTY para conectarse al nodo principal. Antes de conectarse al nodo principal de Amazon EMR, debe descargar e instalar PuTTY y PuTTY. TTYgen Puede descargar estas herramientas desde la página de descarga de PuTTY.

PuTTY no admite de forma nativa el formato de archivo de clave privada (`.pem`) de pares de claves generado por Amazon. EC2 Utiliza Pu TTYgen para convertir el archivo clave al formato PuTTY requerido (`).ppk`. Debe convertir su clave en este formato (`.ppk`) antes de intentar conectarse al nodo principal utilizando PuTTY.

Para obtener más información sobre la conversión de tu clave, consulta Cómo convertir tu clave privada mediante Pu TTYgen en la Guía del EC2 usuario de Amazon.

Para conectarse al nodo principal utilizando PuTTY

1. Asegúrese de permitir el tráfico SSH entrante. Para obtener instrucciones, consulte Antes de conectarse a Amazon EMR: autorice el tráfico entrante.
2. Abra `putty.exe`. También puede lanzar PuTTY desde la lista de programas de Windows.
3. Si es necesario, en la lista Category (Categoría), elija Session (Sesión).
4. En Nombre de host (o dirección IP), escriba `hadoop@`*`MasterPublicDNS`*. Por ejemplo: `hadoop@`*`ec2-###-##-##-###.compute-1.amazonaws.com`*.
5. En la lista Category (Categoría), elija Connection > SSH (Conexión > SSH), Auth (Autenticación).
6. En Private key file for authentication (Archivo de clave privada para la autenticación), elija Browse (Examinar) y seleccione el archivo `.ppk` que ha generado.
7. Selecciona Abrir y, a continuación, Sí para descartar la alerta de seguridad de PuTTY.

> **Important**
>
> Al iniciar sesión en el nodo principal, escriba `hadoop` si se le solicita un nombre de usuario.

8. Cuando haya terminado de trabajar en el nodo principal, puede cerrar la conexión SSH cerrando PuTTY;.

> **Note**
>
> Para evitar que se supere el tiempo de espera de la conexión SSH, puede elegir Connection (Conexión) en la lista Category (Categoría) y seleccionar la opción Enable TCP_keepalives (Habilitar conexiones keepalives de TCP). Si dispone de una sesión de SSH activa en PuTTY, puede cambiar la configuración abriendo el contexto (clic con el botón derecho) en la barra del título de PuTTY y seleccionar Cambiar configuración.

Si tiene dificultades para usar SSH para conectarse a su nodo principal, consulte Solucionar problemas de conexión a su instancia.

Conectarse al nodo principal utilizando la AWS CLI

Puede crear una conexión SSH con el nodo principal mediante Windows y Linux, Unix y Mac OS X. Independientemente de la plataforma, necesitará el nombre de DNS público del nodo principal y la clave privada del par de EC2 claves de Amazon. AWS CLI Si utiliza Linux, Unix o Mac OS X, también debe configurar los permisos en el archivo de clave privada (`.pemo.ppk`), tal y como se muestra en la siguiente. AWS CLI Para configurar los permisos de archivo de clave privada del par de claves

Para conectarse al nodo principal mediante el AWS CLI

1. Asegúrese de permitir el tráfico SSH entrante. Para obtener instrucciones, consulte Antes de conectarse a Amazon EMR: autorice el tráfico entrante.
2. Para recuperar el identificador del clúster, escriba:

   ```
   aws emr list-clusters
   ```

   El resultado muestra los clústeres, incluido el clúster IDs. Tenga en cuenta el ID del clúster al que se está conectando.

   ```
   "Status": {
       "Timeline": {
           "ReadyDateTime": 1408040782.374,
           "CreationDateTime": 1408040501.213
       },
       "State": "WAITING",
       "StateChangeReason": {
   ```

```
        "Message": "Waiting after step completed"
    }
},
"NormalizedInstanceHours": 4,
"Id": "j-2AL4XXXXXX5T9",
"Name": "AWS CLI cluster"
```

3. Escriba el comando siguiente para abrir una conexión SSH al nodo principal. En el ejemplo siguiente, *j-2AL4XXXXXX5T9* sustitúyalo por el ID del clúster y *~/mykeypair.key* sustitúyelo por la ruta completa y el nombre del `.pem` archivo (para Linux, Unix y Mac OS X) o `.ppk` archivo (para Windows). Por ejemplo, `C:\Users\<username>\.ssh\mykeypair.pem`.

```
aws emr ssh --cluster-id j-2AL4XXXXXX5T9 --key-pair-file ~/mykeypair.key
```

4. Cuando termine de trabajar en el nodo principal, cierre la AWS CLI ventana.

   Para más información, consulte Comandos de Amazon EMR en la AWS CLI. Si tiene dificultades para usar SSH para conectarse a su nodo principal, consulte Solucionar problemas de conexión a su instancia.

Puertos de servicio de Amazon EMR

> **Note**
>
> A continuación se muestran las interfaces y los puertos de servicio para los componentes de Amazon EMR. Esta no es una lista completa de los puertos de servicio. Los servicios no predeterminados, como los puertos SSL y los distintos tipos de protocolos, no aparecen en la lista.

> **Important**
>
> Actúe con precaución al editar las reglas de grupos de seguridad para abrir puertos. Asegúrese de agregar reglas que solo permitan el tráfico desde los clientes de confianza y autenticados para los protocolos y puertos necesarios para ejecutar las cargas de trabajo.

Componente	Descripción del servicio	El servicio se ejecuta de forma predeterminada	Puerto	Clave de configuración
Hadoop	API de REST de KMS HTTP	Sí	9600	hadoop.kms.http.port
HDFS	Interfaz de usuario web de namenode	Sí	9870	dfs.namenode.http-address
	RPC de namenode	Sí	8020	dfs.namenode.rpc-address
	DataNode Interfaz de usuario web	Sí	9864	dfs.datanode.http.address
	Datanode HTTP para la transferencia de datos	Sí	9866	dfs.datanode.address
	RPC de datanode para la transferencia de datos	Sí	9867	dfs.datanode.ipc.address
Hive	HiveServer2 Thrift	Sí	10000	hive.server2.thrift.port
	HiveServer2 HTTP	No	10001	hive.server2.thrift.http.port
	HiveServer2. Interfaz de usuario web	Sí	10002	hive.server2.webui.port
	Metaalmacén de Hive	Sí	9083	hive.metastore.port / metastore.thrift.port
	Web HCat	No	50111	templeton.port

Componente	Descripción del servicio	El servicio se ejecuta de forma predeterminada	Puerto	Clave de configuración
	Servicio de administración de daemons (RPC) de LLAP	No	15004	hive.llap.management.rpc.port
	Puerto YARN shuffle para reproducción aleatoria LLAP-daemon-hosted	No	15551	hive.llap.daemon.yarn.shuffle.port
	RPC del daemon de LLAP	No	Dinámico	hive.llap.daemon.rpc.port
	Interfaz web del daemon de LLAP	No	15002	hive.llap.daemon.web.port
	Servicio de salida del daemon de LLAP	No	15003	hive.llap.daemon.output.service.port
Oozie		Sí	11000	
Tez	Tez UI	Sí	8080	
YARN	Mezclas aleatorias	Sí	13562	mapreduce.shuffle.port
	RPC del localizador	Sí	8040	yarn.nodemanager.localizer.address
		Sí	8041	

Componente	Descripción del servicio	El servicio se ejecuta de forma predeterminada	Puerto	Clave de configuración
	Dirección de la aplicación web de NM	Sí	8042	yarn.nodemanager.webapp.address
	Aplicación web de RM	Sí	8088	yarn.resourcemanager.webapp.address
		Sí	8025	
	Programador	Sí	8030	yarn.resourcemanager.scheduler.address
	interfaz del administrador de aplicaciones	Sí	8032	yarn.resourcemanager.address
	Interfaz de administración de RM	Sí	8033	yarn.resourcemanager.admin.address
	JobHistory Interfaz de usuario web del servidor	Sí	1988	mapreduce.jobhistory.webapp.address
	JobHistory Interfaz de usuario web del administrador del servidor	Sí	10033	mapreduce.jobhistory.admin.address

Componente	Descripción del servicio	El servicio se ejecuta de forma predeterminada	Puerto	Clave de configuración
	JobHistory Servidor (RPC)	Sí	10020	mapreduce.jobhistory.address
	Servidor de cronología de aplicaciones (RPC)	Sí	10200	yarn.timeline-service.address
	Interfaz de usuario web HTTP del servidor de cronología de aplicaciones	Sí	8188	yarn.timeline-service.webapp.address
	Interfaz de usuario web HTTPS del servidor de cronología de aplicaciones	No	8190	yarn.timeline-service.webapp.https.address
		Sí	20888	
ZooKeeper	Puerto del cliente	Sí	2181	
		Sí	37301	
		Sí	8341	

Ver las interfaces web alojadas en clústeres de Amazon EMR

> **Important**
>
> Es posible configurar un grupo de seguridad personalizado para permitir el acceso entrante a estas interfaces web. Tenga en cuenta que cualquier puerto en el que permita el tráfico entrante representa una posible vulnerabilidad de seguridad. Revise con atención los grupos de seguridad personalizados para asegurarse de minimizar las vulnerabilidades. Para obtener más información, consulte Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

Hadoop y otras aplicaciones que instale en el clúster de EMR, publican interfaces de usuario como sitios web alojados en el nodo principal. Por motivos de seguridad, cuando se utilizan grupos de seguridad administrados de Amazon EMR, estos sitios web solo están disponibles en el servidor web local del nodo principal. Por ese motivo, debe conectarse al nodo principal para ver las interfaces web. Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH. Hadoop también publica las interfaces de usuario como sitios web alojados en los nodos secundarios y de tareas. Estos sitios web también se encuentran disponibles solo en servidores web locales en los nodos.

En la tabla siguiente, se muestran las interfaces web que puede ver en las instancias de clúster. Estas interfaces de Hadoop están disponibles en todos los clústeres. Para las interfaces de la instancia maestra, *master-public-dns-name* sustitúyalas por el DNS público maestro que aparece en la pestaña Resumen del clúster de la consola de Amazon EMR. Para las interfaces de instancias principales y de tareas, *coretask-public-dns-name* sustitúyalas por el nombre de DNS público que aparece en la lista de la instancia. Para buscar el nombre DNS público de una instancia, vaya a la consola de Amazon EMR, seleccione el clúster en la lista, seleccione la pestaña Hardware, seleccione el ID del grupo de instancias que contiene la instancia a la que desea conectarse y, a continuación, anote el nombre DNS público que aparece para la instancia.

Nombre de interfaz	URI
Servidor del historial de Flink (versión 5.33 y posteriores de EMR)	http: //:8082/ *master-public-dns-name*
Ganglia	http://ganglia/ *master-public-dns-name*

Nombre de interfaz	URI
Hadoop HDFS (versión NameNode EMR anterior a la 6.x)	*master-public-dns-name* https://:50470/
Hadoop HDFS NameNode	http: //:50070/ *master-public-dns-name*
Hadoop HDFS DataNode	http: //:50075/ *coretask-public-dns-name*
Hadoop HDFS (NameNode EMR versión 6.x)	*master-public-dns-name* https://:9870/
Hadoop HDFS (versión DataNode EMR anterior a la 6.x)	*coretask-public-dns-name* https://:50475/
Hadoop HDFS (DataNode EMR versión 6.x)	*coretask-public-dns-name* https://:9865/
HBase	http: //:16010/ *master-public-dns-name*
Hue	http: //:8888/ *master-public-dns-name*
JupyterHub	https://:9443/ *master-public-dns-name*
Livy	http: //:8998/ *master-public-dns-name*
Chispa HistoryServer	http: //:18080/ *master-public-dns-name*
Tez	http://:8080/tez-ui *master-public-dns-name*
HILO NodeManager	http: //:8042/ *coretask-public-dns-name*
HILO ResourceManager	http: //:8088/ *master-public-dns-name*
Zeppelin	http: //:8890/ *master-public-dns-name*

Dado que hay varias interfaces específicas de aplicación disponibles en el nodo principal que no están disponibles en los nodos secundarios y de tarea, las instrucciones en este documento son específicas del nodo principal de Amazon EMR. El acceso a las interfaces web en los nodos

secundarios y de tarea puede realizarse de la misma manera en que se accedería a las interfaces web en el nodo principal.

Existen varias formas en las que puede acceder a las interfaces web en el nodo principal. El método más rápido y más sencillo consiste en utilizar SSH para conectarse al nodo principal y utilizar el navegador basado en texto, Lynx, para ver los sitios web en su cliente SSH. Sin embargo, Lynx es un navegador basado en texto con una interfaz de usuario limitada que no puede mostrar gráficos. El siguiente ejemplo muestra cómo abrir la ResourceManager interfaz de Hadoop mediante Lynx (los Lynx también URLs se proporcionan al iniciar sesión en el nodo principal mediante SSH).

```
lynx http://ip-###-##-##-###.us-west-2.compute.internal:8088/
```

Existen dos opciones restantes para acceder a las interfaces web en el nodo principal que proporcionan una funcionalidad del navegador completa. Seleccione una de las siguientes opciones:

- Opción 1 (recomendada para usuarios más técnicos): utilice un cliente SSH para conectarse al nodo principal, configure el túnel SSH con enrutamiento de puertos local y utilice un navegador de Internet para abrir interfaces web alojadas en el nodo principal. Este método le permite configurar el acceso a la interfaz web sin usar un proxy SOCKS.
- Opción 2 (recomendada para nuevos usuarios): utilice un cliente SSH para conectarse al nodo principal, configure la tunelización SSH con un reenvío dinámico de puertos y configure su navegador de Internet FoxyProxy para que utilice un complemento, como Firefox o Chrome, a fin de administrar la configuración del proxy SOCKS. SwitchyOmega Este método te permite filtrar automáticamente en URLs función de los patrones de texto y limitar la configuración del proxy a los dominios que coincidan con la forma del nombre DNS del nodo principal. Para obtener más información sobre cómo configurar FoxyProxy Firefox y Google Chrome, consulteOpción 2, parte 2: configurar ajustes de proxy para ver sitios web alojados en el nodo principal del clúster de Amazon EMR.

> **Note**
>
> Si modifica el puerto en el que se ejecuta una aplicación mediante la configuración de clúster, el hipervínculo al puerto no se actualizará en la consola de Amazon EMR. Esto se debe a que la consola no tiene la funcionalidad de leer la configuración de `server.port`.

Con Amazon EMR versión 5.25.0 o posterior, puede acceder a la interfaz de usuario del servidor del historial de Spark desde la consola sin configurar un proxy web a través de una conexión SSH. Para obtener más información, consulte Acceso de un clic al servidor del historial de Spark persistente.

Temas

- Opción 1: configurar un túnel SSH al nodo principal de Amazon EMR utilizando el enrutamiento de puertos local
- Opción 2, parte 1: configurar un túnel SSH al nodo principal utilizando el enrutamiento de puertos dinámico
- Opción 2, parte 2: configurar ajustes de proxy para ver sitios web alojados en el nodo principal del clúster de Amazon EMR

Opción 1: configurar un túnel SSH al nodo principal de Amazon EMR utilizando el enrutamiento de puertos local

Para conectarse al servidor web local en el nodo principal, debe crear un túnel SSH entre su equipo y el nodo principal. Esto se conoce como enrutamiento de puertos. Si no desea utilizar un proxy SOCKS, puede configurar un túnel SSH al nodo principal a través del enrutamiento de puertos local. Con el enrutamiento de puertos local, especifique los puertos locales no utilizados que se utilizan para reenviar el tráfico a puertos remotos específicos en el servidor web local del nodo principal.

La configuración de un túnel SSH utilizando el enrutamiento de puerto local requiere el nombre de DNS pública del nodo principal y el archivo de clave privado de par de claves. Para obtener información sobre cómo localizar el nombre de DNS público principal, consulte Recuperar el nombre de DNS público del nodo principal. Para obtener más información sobre cómo acceder a tu par de claves, consulta los pares de EC2 claves de Amazon en la Guía del EC2 usuario de Amazon. Para obtener más información sobre los sitios que podría desear ver en el nodo principal, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

Configurar un túnel SSH al nodo principal utilizando el enrutamiento de puertos local con OpenSSH

Para configurar un túnel SSH utilizando el enrutamiento de puertos local en terminal

1. Asegúrese de permitir el tráfico SSH entrante. Para obtener instrucciones, consulte Antes de conectarse a Amazon EMR: autorice el tráfico entrante.
2. Abra una ventana de terminal. En Mac OS X, elija Aplicaciones > Utilidades > Terminal. En otras distribuciones de Linux, el terminal normalmente se encuentra en Aplicaciones > Accesorios > Terminal.

3. Escriba el siguiente comando para abrir un túnel SSH en su máquina local. Este comando de ejemplo accede a la interfaz ResourceManager web reenviando el tráfico del puerto local 8157 (un puerto local no utilizado elegido al azar) al puerto 8088 del servidor web local del nodo maestro.

   En el comando, *~/mykeypair.pem* sustitúyalo por la ubicación y el nombre de archivo de tu `.pem` archivo y sustitúyelo por *ec2-###-##-##-###.compute-1.amazonaws.com* el nombre de DNS público maestro de tu clúster. Para acceder a una interfaz web diferente, `8088` sustitúyala por el número de puerto correspondiente. Por ejemplo, `8088` sustitúyalo `8890` por para la interfaz de Zeppelin.

   ```
   ssh -i ~/mykeypair.pem -N -L 8157:ec2-###-##-##-
   ###.compute-1.amazonaws.com:8088 hadoop@ec2-###-##-##-###.compute-1.amazonaws.com
   ```

   `-L` hace referencia al uso de enrutamiento de puertos local que le permite especificar un puerto local utilizado para reenviar datos al puerto remoto identificado en el servidor web local del nodo principal.

   Después de emitir este comando, el terminal permanece abierto y no devuelve una respuesta.
4. Para abrir la interfaz ResourceManager web en su navegador, escriba `http://localhost:8157/` en la barra de direcciones.
5. Cuando haya terminado de trabajar con las interfaces web en el nodo principal, cierre las ventanas del terminal.

Opción 2, parte 1: configurar un túnel SSH al nodo principal utilizando el enrutamiento de puertos dinámico

Para conectarse al servidor web local en el nodo principal, debe crear un túnel SSH entre su equipo y el nodo principal. Esto se conoce como enrutamiento de puertos. Si crea el túnel SSH con enrutamiento de puertos dinámico, todo el tráfico dirigido a un puerto local especificado sin utilizar se enruta al servidor web local en el nodo principal. Esto crea un proxy SOCKS. A continuación, puede configurar su navegador de Internet para usar un complemento, por ejemplo, FoxyProxy o SwitchyOmega para administrar la configuración del proxy SOCKS.

El uso de un complemento de administración de proxy le permite filtrar automáticamente en URLs función de los patrones de texto y limitar la configuración del proxy a los dominios que coincidan con la forma del nombre DNS público del nodo principal. El complemento del navegador gestiona

automáticamente la activación y desactivación del proxy al cambiar entre la visualización de sitios web alojados en el nodo principal y en Internet.

Antes de comenzar, necesita el nombre de DNS pública del nodo principal y el archivo de clave privada del par de claves. Para obtener información sobre cómo localizar el nombre de DNS público principal, consulte Recuperar el nombre de DNS público del nodo principal. Para obtener más información sobre cómo acceder a tu par de claves, consulta los pares de EC2 claves de Amazon en la Guía del EC2 usuario de Amazon. Para obtener más información sobre los sitios que podría desear ver en el nodo principal, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

Configurar un túnel SSH al nodo principal utilizando el enrutamiento de puertos dinámico con OpenSSH

Para configurar un túnel SSH utilizando el enrutamiento de puertos dinámico con OpenSSH

1. Asegúrese de permitir el tráfico SSH entrante. Para obtener instrucciones, consulte Antes de conectarse a Amazon EMR: autorice el tráfico entrante.
2. Abra una ventana de terminal. En Mac OS X, elija Aplicaciones > Utilidades > Terminal. En otras distribuciones de Linux, el terminal normalmente se encuentra en Aplicaciones > Accesorios > Terminal.
3. Escriba el siguiente comando para abrir un túnel SSH en su máquina local. *~/mykeypair.pem*Sustitúyalos por la ubicación y el nombre del `.pem` archivo, *8157* sustitúyalos por un número de puerto local no utilizado y *ec2-###-##-##-###.compute-1.amazonaws.com* sustitúyalos por el nombre de DNS público principal de su clúster.

   ```
   ssh -i ~/mykeypair.pem -N -D 8157 hadoop@ec2-###-##-##-###.compute-1.amazonaws.com
   ```

   Después de emitir este comando, el terminal permanece abierto y no devuelve una respuesta.

   > **Note**
   >
   > `-D` hace referencia al uso de enrutamiento de puertos dinámico que le permite especificar un puerto local utilizado para reenviar datos a todos los puertos remotos en el servidor web local del nodo principal. El enrutamiento de puertos dinámicos crea una escucha de proxy SOCKS local en el puerto especificado en el comando.

4. Una vez que el túnel está activo, configure un proxy SOCKS para su navegador. Para obtener más información, consulte Opción 2, parte 2: configurar ajustes de proxy para ver sitios web alojados en el nodo principal del clúster de Amazon EMR.
5. Cuando haya terminado de trabajar con las interfaces web en el nodo principal, cierre la ventana del terminal.

Configure un túnel SSH mediante el reenvío dinámico de puertos con el AWS CLI

Puede crear una conexión SSH con el nodo principal mediante Windows y Linux, Unix y Mac OS X. Si utiliza el nodo principal AWS CLI en Linux, Unix o Mac OS X, debe establecer los permisos en el `.pem` archivo tal y como se muestra en. AWS CLI Para configurar los permisos de archivo de clave privada del par de claves Si está utilizando Windows, PuTTY debe aparecer en la variable de entorno path o podría recibir un error como OpenSSH o PuTTY no disponible. AWS CLI

Para configurar un túnel SSH mediante el reenvío dinámico de puertos con el AWS CLI

1. Asegúrese de permitir el tráfico SSH entrante. Para obtener instrucciones, consulte Antes de conectarse a Amazon EMR: autorice el tráfico entrante.
2. Crear una conexión SSH con el nodo principal, tal y como se muestra en Conectarse al nodo principal utilizando la AWS CLI.
3. Para recuperar el identificador del clúster, escriba:

```
aws emr list-clusters
```

El resultado muestra los clústeres, incluido el clúster. IDs Tenga en cuenta el ID del clúster al que se está conectando.

```
"Status": {
    "Timeline": {
        "ReadyDateTime": 1408040782.374,
        "CreationDateTime": 1408040501.213
    },
    "State": "WAITING",
    "StateChangeReason": {
        "Message": "Waiting after step completed"
    }
},
"NormalizedInstanceHours": 4,
"Id": "j-2AL4XXXXXX5T9",
```

```
"Name": "AWS CLI cluster"
```

4. Escriba el siguiente comando para abrir un túnel SSH en el nodo principal mediante el enrutamiento de puertos dinámico. En el ejemplo siguiente, *j-2AL4XXXXXX5T9* sustitúyalo por el ID del clúster y *~/mykeypair.key* sustitúyelo por la ubicación y el nombre del `.pem` archivo (para Linux, Unix y Mac OS X) o `.ppk` archivo (para Windows).

```
aws emr socks --cluster-id j-2AL4XXXXXX5T9 --key-pair-file ~/mykeypair.key
```

> **Note**
>
> El comando de socks configura automáticamente el enrutamiento de puertos dinámico en el puerto local 8157. Actualmente este ajuste no se puede modificar.

5. Una vez que el túnel está activo, configure un proxy SOCKS para su navegador. Para obtener más información, consulte Opción 2, parte 2: configurar ajustes de proxy para ver sitios web alojados en el nodo principal del clúster de Amazon EMR.
6. Cuando termine de trabajar con las interfaces web del nodo principal, cierre la AWS CLI ventana.

   Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Configurar un túnel SSH en el nodo principal mediante PuTTY

Los usuarios de Windows pueden utilizar un cliente SSH como PuTTY para crear un túnel SSH al nodo principal. Antes de conectarse al nodo principal de Amazon EMR, debe descargar e instalar PuTTY y PuTTY. TTYgen Puede descargar estas herramientas desde la página de descarga de PuTTY.

PuTTY no admite de forma nativa el formato de archivo de clave privada (`.pem`) de pares de claves generado por Amazon. EC2 Utiliza Pu TTYgen para convertir el archivo clave al formato PuTTY requerido (`).ppk`. Debe convertir su clave en este formato (`.ppk`) antes de intentar conectarse al nodo principal utilizando PuTTY.

Para obtener más información sobre la conversión de tu clave, consulta Cómo convertir tu clave privada mediante Pu TTYgen en la Guía del EC2 usuario de Amazon.

Para configurar un túnel SSH utilizando el enrutamiento de puertos dinámico en PuTTY

1. Asegúrese de permitir el tráfico SSH entrante. Para obtener instrucciones, consulte Antes de conectarse a Amazon EMR: autorice el tráfico entrante.
2. Haga doble clic en `putty.exe` para iniciar PuTTY. También puede lanzar PuTTY desde la lista de programas de Windows.

> **Note**
>
> Si ya tiene una sesión de SSH activa con el nodo principal, puede agregar un túnel pulsando con el botón derecho del ratón en la barra del título de PuTTY y seleccionar Cambiar configuración.

3. Si es necesario, en la lista Category (Categoría), elija Session (Sesión).
4. En el campo Nombre de host, escriba **hadoop@***MasterPublicDNS*. Por ejemplo: **hadoop@***ec2-###-##-##-###.compute-1.amazonaws.com*.
5. En la lista Category (Categoría), expanda Connection > SSH (Conexión > SSH) y, a continuación, elija Auth (Autenticación).
6. En Private key file for authentication (Archivo de clave privada para la autenticación), elija Browse (Examinar) y seleccione el archivo `.ppk` que ha generado.

> **Note**
>
> PuTTY no admite de forma nativa el formato de archivo de clave privada (`.pem`) de pares de claves generado por Amazon. EC2 Utiliza Pu TTYgen para convertir el archivo clave al formato PuTTY requerido (`).ppk`. Debe convertir su clave en este formato (`.ppk`) antes de intentar conectarse al nodo principal utilizando PuTTY.

7. En la lista Category (Categoría), expanda Connection > SSH (Conexión > SSH) y, a continuación, elija Tunnels (Túneles).
8. En el campo Puerto de origen, escriba `8157` (un puerto local no utilizado) y, a continuación, seleccione Agregar.
9. Deje el campo Destination (Destino) en blanco.
10. Seleccione las opciones Dynamic (Dinámico) y Auto.
11. Elija Open.
12. Elija Yes (Sí) para descartar la alerta de seguridad de PuTTY.

> **Important**
>
> Cuando inicia sesión en el nodo principal, escriba `hadoop` si se le solicita un nombre de usuario.

13. Una vez que el túnel está activo, configure un proxy SOCKS para su navegador. Para obtener más información, consulte Opción 2, parte 2: configurar ajustes de proxy para ver sitios web alojados en el nodo principal del clúster de Amazon EMR.
14. Cuando haya terminado de trabajar con las interfaces web en el nodo principal, cierre la ventana de PuTTY.

Opción 2, parte 2: configurar ajustes de proxy para ver sitios web alojados en el nodo principal del clúster de Amazon EMR

Si utiliza un túnel de SSH con enrutamiento de puertos dinámico, debe utilizar un complemento de administración de proxy de SOCKS para controlar los ajustes del proxy en el navegador. El uso de una herramienta de administración de proxy SOCKS le permite filtrar automáticamente en URLs función de los patrones de texto y limitar la configuración del proxy a los dominios que coincidan con la forma del nombre DNS público del nodo principal. El complemento del navegador gestiona automáticamente la activación y desactivación del proxy al cambiar entre la visualización de sitios web alojados en el nodo principal y en Internet. Para administrar la configuración del proxy, configura tu navegador para que utilice un complemento como FoxyProxy o SwitchyOmega.

Para obtener más información acerca de la creación de un túnel de SSH, consulte Opción 2, parte 1: configurar un túnel SSH al nodo principal utilizando el enrutamiento de puertos dinámico. Para obtener más información acerca de las interfaces web disponibles, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

Incluya los siguientes ajustes cuando configure su complemento de proxy:

- Use localhost como dirección de host.
- Use el mismo número de puerto local que seleccionó para establecer el túnel SSH con el nodo principal en Opción 2, parte 1: configurar un túnel SSH al nodo principal utilizando el enrutamiento de puertos dinámico. Por ejemplo, puerto*8157*. Este puerto también debe coincidir con el número de puerto que utiliza en PuTTY u en otro emulador de terminal que utilice para conectarse.

- Especifique el protocolo SOCKS v5. SOCKS v5 le permite configurar de forma opcional la autorización de usuario.
- URL Patterns (Patrones de URL)

  Los siguientes patrones de URL deben autorizarse y especificarse con un tipo de patrón comodín:

  - Los patrones *ec2*.*compute*.amazonaws.com* y *10*.amazonaws.com* coinciden con el nombre de DNS público de los clústeres de las regiones de EE. UU.
  - Los patrones *ec2*.compute* y *10*.compute* coinciden con el nombre de DNS público de los clústeres de todas las demás regiones.
  - Un 10. * patrón para proporcionar acceso a los archivos de JobTracker registro de Hadoop. Modifique este filtro si entra en conflicto con su plan de acceso de red.
  - Los patrones *.ec2.internal* y *.compute.internal* hacen coincidir los nombres de DNS privados (internos) de los clústeres de la región `us-east-1` y de todas las demás regiones, respectivamente.

Ejemplo: configurar para Firefox FoxyProxy

El siguiente ejemplo muestra una configuración FoxyProxy estándar (versión 7.5.1) para Mozilla Firefox.

FoxyProxy proporciona un conjunto de herramientas de administración de proxy. Le permite utilizar un servidor proxy para URLs que los patrones de coincidencia correspondientes a los dominios utilizados por las EC2 instancias de Amazon en su clúster de Amazon EMR.

Para instalarlo y configurarlo FoxyProxy con Mozilla Firefox

1. En Firefox, ve a https://addons.mozilla.org/, busca FoxyProxy Estándar y sigue las instrucciones para añadirlo FoxyProxy a Firefox.
2. Mediante un editor de texto, cree un archivo JSON denominado `foxyproxy-settings.json` con la siguiente configuración de ejemplo.

```
{
  "k20d21508277536715": {
    "active": true,
    "address": "localhost",
    "port": 8157,
    "username": "",
    "password": "",
```

```
    "type": 3,
    "proxyDNS": true,
    "title": "emr-socks-proxy",
    "color": "#0055E5",
    "index": 9007199254740991,
    "whitePatterns": [
      {
        "title": "*ec2*.*compute*.amazonaws.com*",
        "active": true,
        "pattern": "*ec2*.*compute*.amazonaws.com*",
        "importedPattern": "*ec2*.*compute*.amazonaws.com*",
        "type": 1,
        "protocols": 1
      },
      {
        "title": "*ec2*.compute*",
        "active": true,
        "pattern": "*ec2*.compute*",
        "importedPattern": "*ec2*.compute*",
        "type": 1,
        "protocols": 1
      },
      {
        "title": "10.*",
        "active": true,
        "pattern": "10.*",
        "importedPattern": "http://10.*",
        "type": 1,
        "protocols": 2
      },
      {
        "title": "*10*.amazonaws.com*",
        "active": true,
        "pattern": "*10*.amazonaws.com*",
        "importedPattern": "*10*.amazonaws.com*",
        "type": 1,
        "protocols": 1
      },
      {
        "title": "*10*.compute*",
        "active": true,
        "pattern": "*10*.compute*",
        "importedPattern": "*10*.compute*",
        "type": 1,
```

```
        "protocols": 1
      },
      {
        "title": "*.compute.internal*",
        "active": true,
        "pattern": "*.compute.internal*",
        "importedPattern": "*.compute.internal*",
        "type": 1,
        "protocols": 1
      },
      {
        "title": "*.ec2.internal* ",
        "active": true,
        "pattern": "*.ec2.internal*",
        "importedPattern": "*.ec2.internal*",
        "type": 1,
        "protocols": 1
      }
    ],
    "blackPatterns": []
  },
  "logging": {
    "size": 100,
    "active": false
  },
  "mode": "patterns",
  "browserVersion": "68.12.0",
  "foxyProxyVersion": "7.5.1",
  "foxyProxyEdition": "standard"
}
```

3. Abra la página Administrar sus extensiones de Firefox (vaya a about:addons y seleccione Extensiones).
4. Selecciona FoxyProxy Estándar y, a continuación, selecciona el botón de más opciones (el botón que tiene forma de puntos suspensivos).
5. Seleccione Opciones en el menú desplegable.
6. Seleccione Importar ajustes en el menú de la izquierda.
7. En la página Configuración de importación, selecciona Configuración de importación en Configuración de importación a partir de la versión FoxyProxy 6.0, busca la ubicación del **`foxyproxy-settings.json`** archivo que has creado, selecciona el archivo y selecciona Abrir.

8. Pulse Aceptar cuando se le pida para sobrescribir los ajustes existentes y guarde la nueva configuración.

Ejemplo: configurar SwitchyOmega para Chrome

El siguiente ejemplo muestra cómo configurar la SwitchyOmega extensión para Google Chrome. SwitchyOmega permite configurar, administrar y cambiar entre varios proxies.

Para instalarlo y configurarlo SwitchyOmega mediante Google Chrome

1. Ve a la https://chrome.google.com/webstore/categoría o extensiones, busca Proxy y SwitchyOmega añádelo a Chrome.
2. Seleccione Nuevo perfil e introduzca `emr-socks-proxy` como nombre del perfil.
3. Seleccione Perfil de PAC y, a continuación, Crear. Los archivos de configuración automática de proxy (PAC) le ayudan a definir una lista de permisos para las solicitudes del navegador que deben reenviarse a un servidor proxy web.
4. En el campo PAC Script, sustituya el contenido por el siguiente script, que define cuáles URLs deben reenviarse a través de su servidor proxy web. Si especificó un número de puerto diferente al configurar el túnel SSH, sustitúyalo por *8157* su número de puerto.

```
function FindProxyForURL(url, host) {
    if (shExpMatch(url, "*ec2*.*compute*.amazonaws.com*")) return 'SOCKS5
 localhost:8157';
    if (shExpMatch(url, "*ec2*.compute*")) return 'SOCKS5 localhost:8157';
    if (shExpMatch(url, "http://10.*")) return 'SOCKS5 localhost:8157';
    if (shExpMatch(url, "*10*.compute*")) return 'SOCKS5 localhost:8157';
    if (shExpMatch(url, "*10*.amazonaws.com*")) return 'SOCKS5 localhost:8157';
    if (shExpMatch(url, "*.compute.internal*")) return 'SOCKS5 localhost:8157';
    if (shExpMatch(url, "*ec2.internal*")) return 'SOCKS5 localhost:8157';
    return 'DIRECT';
}
```

5. En Acciones, seleccione Aplicar cambios para guardar la configuración del proxy.
6. En la barra de herramientas de Chrome, elige SwitchyOmega y selecciona el `emr-socks-proxy` perfil.

Acceder a una interfaz web en el navegador

Para abrir una interfaz web, introduzca el nombre de DNS público de su nodo principal o de núcleo seguido del número de puerto de la interfaz seleccionada en la barra de direcciones del navegador. En el siguiente ejemplo, se muestra la URL que debes introducir para conectarte al Spark HistoryServer.

```
http://master-public-dns-name:18080/
```

Para obtener instrucciones sobre cómo recuperar el nombre de DNS público de un nodo, consulte Recuperar el nombre de DNS público del nodo principal. Para obtener una lista completa de la interfaz web URLs, consulteVer las interfaces web alojadas en clústeres de Amazon EMR.

Envío del trabajo a un clúster de Amazon EMR

En esta sección se describen los métodos que puede utilizar para enviar el trabajo a un clúster de Amazon EMR. Para enviar el trabajo, puede agregar pasos o enviar los trabajos de Hadoop de forma interactiva al nodo principal.

Tenga en cuenta las siguientes reglas de comportamiento de los pasos al enviar los pasos a un clúster:

- El ID de paso puede contener un máximo de 256 caracteres.
- Puede tener hasta 256 pasos pendientes y en ejecución en un clúster.
- Incluso si tiene 256 pasos activos en ejecución en el clúster, puede enviar trabajos de forma interactiva al nodo principal. Puede enviar un número ilimitado de pasos a lo largo de la vida útil de un clúster en ejecución prolongada, pero solo puede haber 256 pasos en ejecución o pendientes en un momento dado.
- Con las versiones 4.8.0 y posteriores de Amazon EMR, excepto la versión 5.0.0, puede cancelar los pasos pendientes. Para obtener más información, consulte Cancelación de pasos al enviar trabajo a un clúster de Amazon EMR.
- Con las versiones 5.28.0 y posteriores de Amazon EMR, puede cancelar los pasos pendientes y en ejecución. También puede optar por ejecutar varios pasos en paralelo para mejorar la utilización del clúster y ahorrar costos. Para obtener más información, consulte Consideraciones para ejecutar varios pasos en paralelo al enviar un trabajo a Amazon EMR.

> **Note**
>
> Para obtener el mejor rendimiento, le recomendamos que almacene las acciones de arranque personalizadas, los scripts y otros archivos que desee utilizar con Amazon EMR en un bucket de Amazon S3 que se encuentre en el Región de AWS mismo lugar que su clúster.

Temas

- Agregar pasos a un clúster con la consola de administración de Amazon EMR
- Añadir pasos a un clúster de Amazon EMR con el AWS CLI
- Consideraciones para ejecutar varios pasos en paralelo al enviar un trabajo a Amazon EMR
- Visualización de los pasos después de enviar el trabajo a un clúster de Amazon EMR
- Cancelación de pasos al enviar trabajo a un clúster de Amazon EMR

Agregar pasos a un clúster con la consola de administración de Amazon EMR

Utilice los siguientes procedimientos para agregar pasos a un clúster con la AWS Management Console. Para obtener información detallada sobre cómo enviar los pasos de aplicaciones de macrodatos específicas, consulte las siguientes secciones de la guía de versiones de Amazon EMR:

- Enviar un paso JAR personalizado
- Enviar un paso de transmisión de Hadoop
- Enviar un paso de Spark
- Enviar un paso de Pig
- Ejecutar un comando o un script como paso
- Transferir los valores a los pasos para ejecutar los scripts de Hive

Agregar pasos durante la creación del clúster

Desde allí AWS Management Console, puede añadir pasos al crear un clúster.

Console

Para agregar pasos al crear un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En Pasos, seleccione Agregar paso. Escriba los valores adecuados en los campos del cuadro de diálogo Agregar paso. Para obtener información sobre cómo dar formato a los argumentos de los pasos, consulte Agregar argumentos de paso. Las opciones varían según el tipo de paso. Para agregar el paso y salir del cuadro de diálogo, seleccione Agregar paso.
4. Elija cualquier otra opción que se aplique a su clúster.
5. Para lanzar el clúster, elija Crear clúster.

Agregar pasos a un clúster en ejecución

Con el AWS Management Console, puede añadir pasos a un clúster con la opción de finalización automática desactivada.

Console

Para agregar pasos a un clúster en ejecución con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. En la página Pasos de la página de detalles del clúster, seleccione Agregar paso. Para clonar un paso existente, seleccione el menú desplegable Acciones y seleccione Clonar paso.
4. Escriba los valores adecuados en los campos del cuadro de diálogo Agregar paso. Las opciones varían según el tipo de paso. Para agregar el paso y salir del cuadro de diálogo, elija Agregar paso.

Modificar el nivel de simultaneidad de pasos en un clúster en ejecución

Con la AWS Management Console, puede modificar el nivel de simultaneidad de los pasos en un clúster en ejecución.

> **Note**
>
> Solo puede ejecutar varios pasos en paralelo con las versiones 5.28.0 y posteriores de Amazon EMR.

Console

Para modificar la simultaneidad de pasos en un clúster en ejecución con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar. El clúster debe estar en ejecución para cambiar su atributo de simultaneidad.
3. En la pestaña Pasos de la página de detalles del clúster, busque la sección Atributos. Seleccione Editar para cambiar la simultaneidad. Escriba un valor comprendido entre 1 y 256.

Agregar argumentos de paso

Si utiliza AWS Management Console para añadir un paso a su clúster, puede especificar los argumentos para ese paso en el campo Argumentos. Debe separar los argumentos con espacios en blanco y rodear los argumentos de cadena que constan de caracteres y espacios en blanco con comillas.

Example : argumentos correctos

Los argumentos del ejemplo siguiente tienen el formato correcto para el argumento de cadena final AWS Management Console, con comillas.

```
bash -c "aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh ."
```

También puede colocar cada argumento en una línea independiente para facilitar la lectura, como se muestra en el siguiente ejemplo.

```
bash
-c
"aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh ."
```

Example : argumentos incorrectos

Los siguientes argumentos de ejemplo tienen un formato incorrecto para la AWS Management Console. Observe que el último argumento de cadena, `aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh .`, contiene espacios en blanco y no está entre comillas.

```
bash -c aws s3 cp s3://amzn-s3-demo-bucket/my-script.sh .
```

Añadir pasos a un clúster de Amazon EMR con el AWS CLI

En los siguientes procedimientos se muestra cómo agregar pasos a un clúster recién creado y a un clúster en ejecución con la AWS CLI. En ambos ejemplos, el subcomando `--steps` se utiliza para agregar pasos al clúster.

Para añadir pasos durante la creación del clúster

- Escriba el siguiente comando para crear un clúster y añadir un paso de Apache Pig. Asegúrate de *`myKey`* sustituirlo por el nombre de tu EC2 key pair de Amazon.

  ```
  aws emr create-cluster --name "Test cluster" \
  --applications Name=Spark \
  --use-default-roles \
  --ec2-attributes KeyName=myKey \
  --instance-groups InstanceGroupType=PRIMARY,InstanceCount=1,InstanceType=m5.xlarge
   InstanceGroupType=CORE,InstanceCount=2,InstanceType=m5.xlarge \
  --steps '[{"Args":["spark-submit","--deploy-mode","cluster","--
  class","org.apache.spark.examples.SparkPi","/usr/lib/spark/examples/jars/spark-
  examples.jar","5"],"Type":"CUSTOM_JAR","ActionOnFailure":"CONTINUE","Jar":"command-
  runner.jar","Properties":"","Name":"Spark application"}]'
  ```

> **Note**
>
> La lista de argumentos cambia en función del tipo de paso.

De forma predeterminada, el nivel de simultaneidad de pasos es 1. Puede establecer el nivel de simultaneidad de pasos con el parámetro `StepConcurrencyLevel` al crear un clúster.

La salida de un identificador de clúster es similar a la siguiente.

```
{
    "ClusterId": "j-2AXXXXXXGAPLF"
}
```

Para añadir un paso a un clúster en ejecución

- Escriba el siguiente comando para añadir un paso a un clúster en ejecución. Sustituya *j-2AXXXXXXGAPLF* por el ID de su propio clúster.

```
aws emr add-steps --cluster-id j-2AXXXXXXGAPLF \
--steps '[{"Args":["spark-submit","--deploy-mode","cluster","--
class","org.apache.spark.examples.SparkPi","/usr/lib/spark/examples/jars/spark-
examples.jar","5"],"Type":"CUSTOM_JAR","ActionOnFailure":"CONTINUE","Jar":"command-
runner.jar","Properties":"","Name":"Spark application"}]'
```

La salida es un identificador de paso similar al siguiente.

```
{
    "StepIds": [
 "s-Y9XXXXXXAPMD"
    ]
}
```

Para modificarlo StepConcurrencyLevel en un clúster en ejecución

1. En un clúster en ejecución, puede modificar `StepConcurrencyLevel` con la API de `ModifyCluster`. Por ejemplo, escriba el siguiente comando para aumentar `StepConcurrencyLevel` a `10`. Sustituya *`j-2AXXXXXXGAPLF`* por el ID de su clúster.

```
aws emr modify-cluster --cluster-id j-2AXXXXXXGAPLF --step-concurrency-level 10
```

2. El resultado es similar al siguiente.

```
{
"StepConcurrencyLevel": 10
}
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte la Referencia de AWS CLI comandos.

Consideraciones para ejecutar varios pasos en paralelo al enviar un trabajo a Amazon EMR

La ejecución de varios pasos en paralelo al enviar un trabajo a Amazon EMR requiere decisiones preliminares sobre la planificación de los recursos y las expectativas en relación con el comportamiento del clúster. A continuación, se analizan en detalle.

- Los pasos que se ejecutan en paralelo pueden completarse en cualquier orden, pero los pasos pendientes de la cola pasan al estado de ejecución en el orden en que se enviaron.
- Al seleccionar un nivel de simultaneidad de pasos para el clúster, debe considerar si el tipo de instancia del nodo principal cumple o no los requisitos de memoria de las cargas de trabajo de usuario. El proceso de ejecución del paso principal se ejecuta en el nodo principal para cada paso. Para ejecutar varios pasos en paralelo, se requiere más memoria y utilización de CPU en el nodo principal que si los pasos se ejecutan de uno en uno.
- Para gestionar los recursos y realizar las programaciones complejas de los pasos simultáneos, puede utilizar las funciones de programación de YARN como `FairScheduler` o `CapacityScheduler`. Por ejemplo, puede utilizar `FairScheduler` con un conjunto `queueMaxAppsDefault` para evitar que se ejecuten más de un determinado número de trabajos a la vez.

- El nivel de simultaneidad de pasos está sujeto a las configuraciones de los administradores de recursos. Por ejemplo, si YARN está configurado con un paralelismo de solo `5`, entonces únicamente podrá tener cinco aplicaciones YARN ejecutándose en paralelo, aunque `StepConcurrencyLevel` se haya establecido en `10`. Para obtener más información sobre la configuración de los administradores de recursos, consulte Configuración de aplicaciones en la Guía de publicación de Amazon EMR.
- No puede agregar un paso con un valor de `ActionOnFailure` que no sea CONTINUE mientras el nivel de simultaneidad de pasos del clúster sea superior a 1.
- Si el nivel de simultaneidad de pasos de un clúster es superior a uno, la característica `ActionOnFailure` del paso no se activará.
- Si un clúster tiene un nivel de simultaneidad de pasos de `1`, pero tiene varios pasos en ejecución, es posible que `TERMINATE_CLUSTER ActionOnFailure` se active, pero `CANCEL_AND_WAIT ActionOnFailure` no se activará. Este caso de periferia se presenta cuando el nivel de simultaneidad de los pasos del clúster era superior a uno, pero se reducía cuando se ejecutaban varios pasos.
- Puede usar el escalado automático de EMR para escalar en sentido ascendente o descendente en función de los recursos de YARN y así evitar la contención de recursos. Para obtener información, consulte Uso del escalado automático con una política personalizada para grupos de instancias en la Guía de administración de Amazon EMR.
- Cuando se disminuye el nivel de simultaneidad de pasos, EMR permite que se completen todos los pasos en ejecución antes de reducir el número de pasos. Si los recursos se agotan porque el clúster está ejecutando demasiados pasos simultáneos, recomendamos cancelar manualmente los pasos en ejecución para liberar recursos.

Visualización de los pasos después de enviar el trabajo a un clúster de Amazon EMR

Puede ver hasta 10 000 pasos que Amazon EMR haya completado en los últimos siete días. También puede ver los 1000 pasos que Amazon EMR haya completado en cualquier momento. Este total incluye tanto los pasos del sistema como los pasos enviados por el usuario.

Si envía nuevos pasos una vez que el clúster alcanza el límite de registro de 1000 pasos, Amazon EMR eliminará los pasos inactivos enviados por el usuario cuyos estados hayan sido COMPLETADOS, CANCELADOS o FALLIDOS durante más de siete días. Si envía pasos que superen el límite de registros de 10 000 pasos, Amazon EMR eliminará los registros de pasos

inactivos enviados por el usuario, independientemente de su duración inactiva. Amazon EMR no elimina estos registros de los archivos de registro. Amazon EMR los elimina de la AWS consola y no se devuelven cuando se utiliza la API AWS CLI o para recuperar la información del clúster. Los registros de paso del sistema no se eliminan nunca.

La información de paso que puede ver depende del mecanismo utilizado para recuperar la información del clúster. La tabla siguiente indica la información de paso devuelta por cada una de las opciones disponibles.

Opción	DescribeJobFlow o --describe --jobflow	ListSteps o list-steps
SDK	256 pasos	Hasta 10 000 pasos
CLI de Amazon EMR	256 pasos	N/A
AWS CLI	N/A	Hasta 10 000 pasos
API	256 pasos	Hasta 10 000 pasos

Cancelación de pasos al enviar trabajo a un clúster de Amazon EMR

Puede cancelar los pasos pendientes y en ejecución de Amazon EMR al enviar el trabajo a su clúster. AWS Management Console AWS CLI API.

Console

Para cancelar los pasos con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, seleccione el clúster que desee actualizar.
3. En la pestaña Pasos de la página de detalles del clúster, active la casilla de verificación situada junto al paso que desee cancelar. Seleccione el menú desplegable Acciones y, a continuación, seleccione Cancelar pasos.

4. En el cuadro de diálogo Cancelar el paso, seleccione si desea cancelar el paso y esperar a que se cierre o cancelar el paso y forzar el cierre. A continuación, seleccione Confirm (Confirmar).
5. El estado de los pasos de la tabla Pasos cambia a `CANCELLED`.

CLI

Para cancelar con el uso del AWS CLI

- Utilice el comando `aws emr cancel-steps`, especificando el clúster y los pasos que desea cancelar. El siguiente ejemplo muestra un comando de AWS CLI para cancelar dos pasos.

```
aws emr cancel-steps --cluster-id j-2QUAXXXXXXXXX \
--step-ids s-3M8DXXXXXXXXX s-3M8DXXXXXXXXX \
--step-cancellation-option SEND_INTERRUPT
```

Con la versión 5.28.0 de Amazon EMR, puede seleccionar una de las dos siguientes opciones de cancelación para el parámetro `StepCancellationOption` al cancelar pasos.

- `SEND_INTERRUPT`: esta es la opción predeterminada. Cuando se recibe una solicitud de cancelación de pasos, EMR envía una señal `SIGTERM` al paso. Agregue un controlador de señales `SIGTERM` a su lógica de pasos para atrapar esta señal y terminar los procesos de pasos descendientes o espere a que se completen.
- `TERMINATE_PROCESS`: cuando se selecciona esta opción, EMR envía una señal `SIGKILL` al paso y a todos sus procesos descendientes, que los termina inmediatamente.

Consideraciones para cancelar los pasos

- Al cancelar un paso pendiente o en ejecución, ese paso se elimina del recuento de pasos activos.
- Si se cancela un paso en ejecución, no se permite que un paso pendiente comience a ejecutarse, suponiendo que no haya cambios en `stepConcurrencyLevel`.
- La cancelación de un paso en ejecución no activa el paso `ActionOnFailure`.
- Para las versiones 5.32.0 y posteriores de EMR, `SEND_INTERRUPT StepCancellationOption` envía una señal `SIGTERM` al proceso secundario del paso. Debe estar atento a esta señal y realizar una limpieza para apagarlo correctamente. `TERMINATE_PROCESS`

`StepCancellationOption` envía una señal `SIGKILL` al proceso secundario del paso y a todos sus procesos descendientes; sin embargo, los procesos asíncronos no se ven afectados.

Visualización y monitorización de un clúster de Amazon EMR a medida que realiza su trabajo

Amazon EMR proporciona varias herramientas que puede utilizar para recopilar información sobre su clúster. Puede acceder a información sobre el clúster desde la consola, la CLI o mediante programación. Las interfaces web estándar de Hadoop y los archivos de registro están disponibles en el nodo principal. También puede utilizar servicios de supervisión como CloudWatch Ganglia para realizar un seguimiento del rendimiento de su clúster.

El historial de aplicaciones también está disponible desde la consola mediante la aplicación «persistente» UIs para Spark History Server a partir de Amazon EMR 5.25.0. Con Amazon EMR 6.x, el servidor de YARN Timeline persistente y las interfaces de usuario de Tez también están disponibles. Estos servicios están alojados fuera del clúster, por lo que puede acceder al historial de aplicaciones durante 30 días después de que termine el clúster, sin necesidad de una conexión SSH o proxy web. Consulte Ver el historial de aplicaciones.

Temas

- Visualización del estado y los detalles del clúster de Amazon EMR
- Depuración de pasos mejorada con Amazon EMR
- Visualización del historial de aplicaciones de Amazon EMR
- Visualización de los archivos de registro de Amazon EMR
- Ver instancias de clúster en Amazon EC2
- CloudWatch eventos y métricas de Amazon EMR
- Visualización de métricas de aplicaciones de clúster mediante Ganglia con Amazon EMR
- Registro de llamadas a la API AWS EMR mediante AWS CloudTrail
- Mejores prácticas de observabilidad de EMR

Visualización del estado y los detalles del clúster de Amazon EMR

Después de crear un clúster, puede monitorizar su estado y obtener información detallada acerca de su ejecución y los errores que puedan haberse producido, incluso después de que se haya

terminado. Amazon EMR guarda los metadatos de los clústeres terminados para su referencia durante dos meses, después de los cuales se eliminan. No puede eliminar clústeres del historial de clústeres, pero con la AWS Management Console, puede utilizar la función Filter (Filtro), y con la AWS CLI, puede utilizar opciones con el comando `list-clusters` para centrarse en los clústeres que le interesen.

Puede acceder al historial de aplicaciones almacenado en el clúster durante una semana desde el momento en que se registra, independientemente de si el clúster se está ejecutando o ha terminado. Además, las interfaces de usuario de aplicaciones persistentes almacenan el historial de aplicaciones fuera del clúster durante 30 días después de que termine un clúster. Consulte Ver el historial de aplicaciones.

Para obtener más información sobre los estados de los clústeres, como los estados En espera y En ejecución, consulte Descripción del ciclo de vida del clúster.

Ver los detalles de un clúster mediante la AWS Management Console

La lista de clústeres del archivo https://console.aws.amazon.com/emr incluye todos los clústeres de tu cuenta y AWS región, incluidos los clústeres terminados. La lista muestra lo siguiente para cada clúster: el nombre y el identificador, el estado y los detalles del estado, la hora de creación, el tiempo transcurrido durante el que el clúster estuvo en ejecución y las horas de instancia normalizadas acumuladas en todas las instancias del clúster. EC2 Esta lista es el punto de partida para monitorear el estado de los clústeres. Se ha diseñado para que pueda profundizar hasta los detalles de cada clúster para su análisis y resolución de problemas.

Console

Para ver la información del clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee ver.
3. Utilice el panel Resumen para ver los aspectos básicos de la configuración del clúster, como el estado del clúster, las aplicaciones de código abierto que Amazon EMR instaló en el clúster y la versión de Amazon EMR que utilizó para crear el clúster. Utilice les pestañas que aparecen debajo del resumen para ver la información que se describe en la siguiente tabla.

Para ver los detalles del clúster, utilice la AWS CLI

Los siguientes ejemplos muestran cómo recuperar detalles del clúster utilizando la AWS CLI. Para obtener más información sobre los comandos disponibles, consulte la Referencia de comandos de la AWS CLI de Amazon EMR. Puede usar el comando describe-cluster para ver detalles de nivel de clúster como el estado, la configuración de hardware y de software, los ajustes de VPC, las acciones de arranque, los grupos de instancias, etc. Para obtener más información acerca de los estados del clúster, consulte Descripción del ciclo de vida del clúster. El siguiente ejemplo ilustra cómo usar el comando `describe-cluster`, seguido de ejemplos del comando list-clusters.

Example Ver el estado del clúster

Para utilizar el comando `describe-cluster`, necesita el ID del clúster. En este ejemplo, se muestra cómo obtener una lista de los clústeres creados dentro de un intervalo de fechas determinado y, a continuación, utilizar uno de los clústeres IDs devueltos para obtener más información sobre el estado de un clúster individual.

El siguiente comando describe el clúster*j-1K48XXXXXXHCB*, que se reemplaza por el ID del clúster.

```
aws emr describe-cluster --cluster-id j-1K48XXXXXXHCB
```

La salida de este comando es similar a la siguiente:

```
{
    "Cluster": {
        "Status": {
            "Timeline": {
                "ReadyDateTime": 1438281058.061,
                "CreationDateTime": 1438280702.498
            },
            "State": "WAITING",
            "StateChangeReason": {
                "Message": "Waiting for steps to run"
            }
        },
        "Ec2InstanceAttributes": {
            "EmrManagedMasterSecurityGroup": "sg-cXXXXX0",
            "IamInstanceProfile": "EMR_EC2_DefaultRole",
            "Ec2KeyName": "myKey",
            "Ec2AvailabilityZone": "us-east-1c",
            "EmrManagedSlaveSecurityGroup": "sg-example"
        },
```

```
        "Name": "Development Cluster",
        "ServiceRole": "EMR_DefaultRole",
        "Tags": [],
        "TerminationProtected": false,
        "ReleaseLabel": "emr-4.0.0",
        "NormalizedInstanceHours": 16,
        "InstanceGroups": [
            {
                "RequestedInstanceCount": 1,
                "Status": {
                    "Timeline": {
                        "ReadyDateTime": 1438281058.101,
                        "CreationDateTime": 1438280702.499
                    },
                    "State": "RUNNING",
                    "StateChangeReason": {
                        "Message": ""
                    }
                },
                "Name": "CORE",
                "InstanceGroupType": "CORE",
                "Id": "ig-2EEXAMPLEXXP",
                "Configurations": [],
                "InstanceType": "m5.xlarge",
                "Market": "ON_DEMAND",
                "RunningInstanceCount": 1
            },
            {
                "RequestedInstanceCount": 1,
                "Status": {
                    "Timeline": {
                        "ReadyDateTime": 1438281023.879,
                        "CreationDateTime": 1438280702.499
                    },
                    "State": "RUNNING",
                    "StateChangeReason": {
                        "Message": ""
                    }
                },
                "Name": "MASTER",
                "InstanceGroupType": "MASTER",
                "Id": "ig-2A1234567XP",
                "Configurations": [],
                "InstanceType": "m5.xlarge",
```

```
                "Market": "ON_DEMAND",
                "RunningInstanceCount": 1
            }
        ],
        "Applications": [
            {
                "Version": "1.0.0",
                "Name": "Hive"
            },
            {
                "Version": "2.6.0",
                "Name": "Hadoop"
            },
            {
                "Version": "0.14.0",
                "Name": "Pig"
            },
            {
                "Version": "1.4.1",
                "Name": "Spark"
            }
        ],
        "BootstrapActions": [],
        "MasterPublicDnsName": "ec2-X-X-X-X.compute-1.amazonaws.com",
        "AutoTerminate": false,
        "Id": "j-jobFlowID",
        "Configurations": [
            {
                "Properties": {
                    "hadoop.security.groups.cache.secs": "250"
                },
                "Classification": "core-site"
            },
            {
                "Properties": {
                    "mapreduce.tasktracker.reduce.tasks.maximum": "5",
                    "mapred.tasktracker.map.tasks.maximum": "2",
                    "mapreduce.map.sort.spill.percent": "90"
                },
                "Classification": "mapred-site"
            },
            {
                "Properties": {
                    "hive.join.emit.interval": "1000",
```

```
                    "hive.merge.mapfiles": "true"
                },
                "Classification": "hive-site"
            }
        ]
    }
}
```

Example Mostrar clústeres por fecha de creación

Para recuperar clústeres creados en un intervalo de fechas específico, use el comando `list-clusters` con los parámetros `--created-after` y `--created-before`.

El siguiente comando muestra todos los clústeres creados entre el 9 de octubre de 2019 y el 12 de octubre de 2019.

```
aws emr list-clusters --created-after 2019-10-09T00:12:00 --created-
before 2019-10-12T00:12:00
```

Example Mostrar clústeres por estado

Para mostrar clústeres por estado, use el comando `list-clusters` con el parámetro `--cluster-states`. Los estados de clúster válidos incluyen: STARTING, BOOTSTRAPPING, RUNNING, WAITING, TERMINATING, TERMINATED y TERMINATED_WITH_ERRORS.

```
aws emr list-clusters --cluster-states TERMINATED
```

También puede utilizar los siguientes parámetros de acceso directo para mostrar todos los clústeres con los estados especificados:

- El parámetro `--active` filtra los clústeres por los estados STARTING, BOOTSTRAPPING, RUNNING, WAITING o TERMINATING.
- El parámetro `--terminated` filtra los clústeres por el estado TERMINATED.
- El parámetro `--failed` filtra los clústeres por el estado TERMINATED_WITH_ERRORS.

Los siguientes comandos devuelven el mismo resultado.

```
aws emr list-clusters --cluster-states TERMINATED
```

```
aws emr list-clusters --terminated
```

Para obtener más información acerca de los estados del clúster, consulte Descripción del ciclo de vida del clúster.

Depuración de pasos mejorada con Amazon EMR

Si un paso de Amazon EMR devuelve un error y envió el trabajo utilizando la operación de la API de pasos con la versión 5.x o posterior de la AMI, Amazon EMR puede identificar y devolver la causa raíz del error del paso en algunos casos, junto con el nombre del archivo de registro relevante y una parte del registro de seguimiento de la pila de la aplicación a través de la API. Por ejemplo, puede identificar los siguientes errores:

- Un error de Hadoop común como, por ejemplo, el directorio de salida ya existe, el directorio de entrada no existe o una aplicación se queda sin memoria.
- Errores de Java como, por ejemplo, una aplicación que se ha compilado con una versión incompatible de Java o se ha ejecutado con una clase principal que no se encuentra.
- Un problema al acceder a objetos almacenados en Amazon S3.

Esta información está disponible mediante las operaciones DescribeStepy ListStepsAPI. El FailureDetailscampo del StepSummarydevuelto por esas operaciones. Para acceder a la FailureDetails información, utilice la AWS CLI, la consola o el AWS SDK.

Console

Tenga en cuenta que la nueva consola de Amazon EMR no ofrece la depuración de pasos. Sin embargo, puede ver los detalles de la terminación del clúster siguiendo estos pasos.

Para ver los detalles del error con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, seleccione el clúster que desee ver.
3. Anote el valor del estado de la sección Resumen de la página de detalles del clúster. Si el estado es terminado con errores, pase el ratón sobre el texto para ver los detalles de los errores del clúster.

CLI

Para ver los detalles de los fallos con el AWS CLI

- Para obtener los detalles del error de un paso con el AWS CLI, utilice el `describe-step` comando.

```
aws emr describe-step --cluster-id j-1K48XXXXXHCB --step-id s-3QM0XXXXXM1W
```

El resultado tendrá un aspecto similar al siguiente:

```
{
  "Step": {
    "Status": {
      "FailureDetails": {
        "LogFile": "s3://amzn-s3-demo-bucket/logs/j-1K48XXXXXHCB/steps/
s-3QM0XXXXXM1W/stderr.gz",
        "Message": "org.apache.hadoop.mapred.FileAlreadyExistsException: Output
 directory s3://amzn-s3-demo-bucket/logs/beta already exists",
        "Reason": "Output directory already exists."
      },
      "Timeline": {
        "EndDateTime": 1469034209.143,
        "CreationDateTime": 1469033847.105,
        "StartDateTime": 1469034202.881
      },
      "State": "FAILED",
      "StateChangeReason": {}
    },
    "Config": {
      "Args": [
        "wordcount",
        "s3://amzn-s3-demo-bucket/input/input.txt",
        "s3://amzn-s3-demo-bucket/logs/beta"
      ],
      "Jar": "s3://amzn-s3-demo-bucket/jars/hadoop-mapreduce-examples-2.7.2-
amzn-1.jar",
      "Properties": {}
    },
    "Id": "s-3QM0XXXXXM1W",
    "ActionOnFailure": "CONTINUE",
    "Name": "ExampleJob"
```

```
  }
}
```

Visualización del historial de aplicaciones de Amazon EMR

Puede ver los detalles de aplicación del servidor del historial de Spark y de YARN Timeline Service en la página de detalles del clúster en la consola. El historial de aplicaciones de Amazon EMR le permite solucionar problemas y analizar más fácilmente los trabajos activos y el historial de trabajos.

> **Note**
>
> Para aumentar la seguridad de las aplicaciones fuera de la consola que podría utilizar con Amazon EMR, los dominios de alojamiento de aplicaciones se registran en la lista de sufijos públicos (PSL). Algunos ejemplos de estos dominios de alojamiento son los siguientes: `emrstudio-prod.us-east-1.amazonaws.com`, `emrnotebooks-prod.us-east-1.amazonaws.com`, `emrappui-prod.us-east-1.amazonaws.com`. Para mayor seguridad, si necesita configurar cookies confidenciales en el nombre de dominio predeterminado, le recomendamos que utilice cookies con el prefijo `__Host-`. Esta práctica lo ayuda a proteger su dominio de los intentos de falsificación de solicitudes entre sitios (CSRF). Para obtener más información, consulte la página Set-Cookie en la Red de desarrolladores de Mozilla.

La sección Interfaces de usuario de aplicaciones de la pestaña Aplicaciones ofrece varias opciones de visualización, en función del estado del clúster y de las aplicaciones que haya instalado en él.

- Acceso fuera del clúster a las interfaces de usuario de aplicaciones persistentes: a partir de la versión 5.25.0 de Amazon EMR, los enlaces a las interfaces de usuario de aplicaciones persistentes están disponibles para la interfaz de usuario de Spark y el servicio de historial de Spark. Con las versiones 5.30.1 y posteriores de Amazon EMR, Tez UI y el servidor de YARN Timeline también tienen interfaces de usuario de aplicaciones persistentes. El servidor de YARN Timeline y Tez UI son aplicaciones de código abierto que proporcionan métricas y herramientas visuales para clústeres activos y terminados. La interfaz de usuario de Spark ofrece detalles acerca de las etapas y tareas del programador, tamaños de RDD y uso de memoria, información medioambiental e información acerca de los ejecutores en ejecución. UIs Las aplicaciones persistentes se ejecutan fuera del clúster, por lo que la información y los registros del clúster están disponibles durante 30 días después de la finalización de la aplicación. A diferencia de las

interfaces de usuario de las aplicaciones del clúster, las aplicaciones persistentes UIs no requieren que configure un proxy web a través de una conexión SSH.

- Interfaces de usuario de aplicaciones en el clúster: hay una variedad de interfaces de usuario del historial de aplicaciones que se pueden ejecutar en un clúster. Las interfaces de usuario en el clúster están alojadas en el nodo maestro y requieren que configure una conexión SSH al servidor web. Las interfaces de usuario de aplicaciones en el clúster mantienen el historial de aplicaciones durante una semana después de que termine una aplicación. Para obtener más información e instrucciones sobre cómo configurar un túnel SSH, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

  Con la excepción del servidor del historial de Spark, el servidor de YARN Timeline y las aplicaciones de Hive, el historial de aplicaciones en el clúster solo se puede consultar mientras el clúster está en ejecución.

Visualización de interfaces de usuario de aplicaciones persistentes en Amazon EMR

A partir de la versión 5.25.0 de Amazon EMR, puede conectarse a los detalles persistentes de la aplicación del servidor del historial de Spark alojados fuera del clúster mediante la página Resumen del clúster o la pestaña Interfaces de usuario de aplicaciones de la consola. Las interfaces de aplicación persistentes de Tez UI y YARN Timeline están disponibles a partir de la versión 5.30.1 de Amazon EMR. El acceso de enlace con un solo clic al historial de aplicaciones persistentes ofrece las siguientes ventajas:

- Puede analizar y solucionar rápidamente los trabajos activos y el historial de trabajos sin configurar un proxy web a través de una conexión SSH.
- Puede acceder al historial de aplicaciones y a los archivos de registro pertinentes para los clústeres activos y terminados. Los registros están disponibles durante 30 días desde que finalice la aplicación.

Navegue hasta los detalles del clúster en la consola y seleccione la pestaña Aplicaciones. Seleccione la interfaz de usuario de la aplicación que desee una vez que se haya lanzado el clúster. La interfaz de usuario de la aplicación se abre en una nueva pestaña del navegador. Para obtener más información, consulte Monitoring and Instrumentation.

Puede ver los registros de contenedores de YARN a través de los enlace en Spark History Server, el servidor de YARN Timeline y Tez UI.

> **Note**
>
> Para obtener acceso a los registros de contenedor de YARN desde el servidor del historial de Spark, el servidor de YARN Timeline y Tez UI, debe habilitar el registro en Amazon S3 para su clúster. Si el registro no está habilitado, los enlaces a los registros de contenedor de YARN no funcionarán.

Recopilación de registros

Para habilitar el acceso con un solo clic a las interfaces de usuario de aplicaciones persistentes, Amazon EMR recopila dos tipos de registros:

- Los registros de eventos de aplicación se recopilan en un bucket del sistema EMR. Los registros de eventos se cifran en reposo mediante el cifrado del lado del servidor con claves administradas de Amazon S3 (SSE-S3). Si utiliza una subred privada para su clúster, asegúrese de incluir el bucket de sistema correcto ARNs en la lista de recursos de la política de Amazon S3 para la subred privada. Para obtener más información, consulte Política de Amazon S3 mínima para subred privada.
- Los registros de contenedores de YARN se recopilan en un bucket de Amazon S3 propiedad de usted. Debe habilitar el registro para que su clúster acceda a los registros de contenedores de YARN. Para más información, consulte Configurar el registro y la depuración de un clúster.

Si necesita deshabilitar esta característica por motivos de privacidad, puede detener el daemon mediante un script de arranque al crear un clúster, como se muestra en el siguiente ejemplo.

```
aws emr create-cluster --name "Stop Application UI Support" --release-label emr-7.10.0
 \
--applications Name=Hadoop Name=Spark --ec2-attributes KeyName=<myEMRKeyPairName> \
--instance-groups InstanceGroupType=MASTER,InstanceCount=1,InstanceType=m3.xlarge
 InstanceGroupType=CORE,InstanceCount=1,InstanceType=m3.xlarge
 InstanceGroupType=TASK,InstanceCount=1,InstanceType=m3.xlarge \
--use-default-roles --bootstrap-actions Path=s3://region.elasticmapreduce/bootstrap-
actions/run-if,Args=["instance.isMaster=true","echo Stop Application UI | sudo tee /
etc/apppusher/run-apppusher; sudo systemctl stop apppusher || exit 0"]
```

Después de ejecutar este script de arranque, Amazon EMR no recopilará ningún registro de eventos del servidor del historial de Spark ni del servidor de YARN Timeline en el bucket del sistema de EMR.

No habrá información del historial de aplicaciones disponible en la pestaña Interfaces de usuario de aplicaciones y perderá el acceso a todas las interfaces de usuario de la aplicación desde la consola.

Archivos de registro de eventos de Spark de gran tamaño

En algunos casos, los trabajos de Spark de larga duración, como las transmisiones de Spark, y los trabajos de gran tamaño, como las consultas SQL de Spark, pueden generar registros de eventos de gran tamaño. Con registros de eventos de gran tamaño, puede consumir rápidamente espacio en disco en las instancias de cómputo y detectar `OutOfMemory` errores al cargar Persistent. UIs Para evitar estos problemas, recomendamos que active la característica de acumulación y compactación de registros de eventos de Spark. Esta característica solo está disponible en las versiones emr-6.1.0 y posteriores de Amazon EMR. Para obtener más información sobre la acumulación y la compactación, consulte Applying compaction on rolling event log files en la documentación de Spark.

Para activar la característica de acumulación y compactación de registros de eventos de Spark, active los siguientes ajustes en la configuración de Spark.

- `spark.eventLog.rolling.enabled`: activa la acumulación de registros de eventos en función del tamaño. Este ajuste está desactivado de forma predeterminada.
- `spark.eventLog.rolling.maxFileSize`: cuando se activa la acumulación, especifica el tamaño máximo del archivo de registros de eventos antes de que se acumule. El valor predeterminado es 128 MB.
- `spark.history.fs.eventLog.rolling.maxFilesToRetain`: especifica el número máximo de archivos de registros de eventos no compactados que se deben retener. De forma predeterminada, se retienen todos los archivos de registros de eventos. Configúrelo en un número inferior para compactar los registros de eventos más antiguos. El valor más bajo es 1.

Tenga en cuenta que la compactación intenta excluir los eventos con archivos de registros de eventos desactualizados, como los siguientes. Si descarta los eventos, dejará de verlos en la interfaz de usuario del servidor del historial de Spark.

- Eventos para trabajos terminados y eventos relacionados de etapas o tareas.
- Eventos para ejecutores terminados.
- Eventos para consultas SQL completadas y eventos relacionados de trabajos, etapas y tareas.

Para lanzar un clúster con la acumulación y la compactación activadas

1. Cree un archivo `spark-configuration.json` con la siguiente configuración.

```
[
   {
     "Classification": "spark-defaults",
        "Properties": {
           "spark.eventLog.rolling.enabled": true,
           "spark.history.fs.eventLog.rolling.maxFilesToRetain": 1
        }
   }
]
```

2. Cree su clúster con la configuración de compactación de acumulación de Spark de la siguiente manera.

```
aws emr create-cluster \
--release-label emr-6.6.0 \
--instance-type m4.large \
--instance-count 2 \
--use-default-roles \
--configurations file://spark-configuration.json
```

Permisos para ver las interfaces de usuario de las aplicaciones persistentes

En el siguiente ejemplo, se muestran los permisos de rol necesarios para acceder a las interfaces de usuario de las aplicaciones persistentes. En el caso de los clústeres con la función de tiempo de ejecución habilitada, esto solo permitirá a los usuarios acceder a las aplicaciones enviadas por la misma identidad de usuario y función de ejecución.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:CreatePersistentAppUI",
```

```
        "elasticmapreduce:DescribePersistentAppUI"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:cluster/clusterId"
      ],
      "Sid": "AllowELASTICMAPREDUCECreatepersistentappui"
    },
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:GetPersistentAppUIPresignedURL"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:*:123456789012:cluster/clusterId",
        "arn:aws:elasticmapreduce:*:123456789012:persistent-app-ui/*"
      ],
      "Condition": {
        "StringEqualsIfExists": {
          "elasticmapreduce:ExecutionRoleArn": [
            "arn:aws:iam::123456789012:role/executionRoleArn"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEGetpersistentappuipresignedurl"
    }
  ]
}
```

En el siguiente ejemplo, se muestran los permisos de rol necesarios para eliminar las restricciones a la hora de ver las aplicaciones en las interfaces de usuario de las aplicaciones persistentes para los clústeres con funciones en tiempo de ejecución.

JSON

```
{
  "Version": "2012-10-17",
  "Statement": [
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:CreatePersistentAppUI",
```

```
        "elasticmapreduce:DescribePersistentAppUI",
        "elasticmapreduce:AccessAllEventLogs"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:us-east-1:123456789012:cluster/j-XXXXXXXXXXXXX"
      ],
      "Sid": "AllowELASTICMAPREDUCECreatepersistentappui"
    },
    {
      "Effect": "Allow",
      "Action": [
        "elasticmapreduce:GetPersistentAppUIPresignedURL"
      ],
      "Resource": [
        "arn:aws:elasticmapreduce:us-east-1:123456789012:cluster/j-
XXXXXXXXXXXXX",
        "arn:aws:elasticmapreduce:us-east-1:123456789012:persistent-app-ui/*"
      ],
      "Condition": {
        "StringEqualsIfExists": {
          "elasticmapreduce:ExecutionRoleArn": [
            "arn:aws:iam::123456789012:role/YourExecutionRoleName"
          ]
        }
      },
      "Sid": "AllowELASTICMAPREDUCEGetpersistentappuipresignedurl"
    }
  ]
}
```

Consideraciones y limitaciones

El acceso con un solo clic a las interfaces de usuario de aplicaciones persistentes actualmente tiene las siguientes limitaciones.

- Habrá un retraso de al menos dos minutos cuando los detalles de la aplicación aparezcan en la interfaz de usuario del servidor del historial de Spark.
- Esta característica solo funciona cuando el directorio de registros de eventos de la aplicación está en HDFS. De forma predeterminada, Amazon EMR almacena los registros de eventos en un directorio de HDFS. Si cambia el directorio predeterminado a un sistema de archivos diferente, como Amazon S3, esta característica no funcionará.

- Esta función no está disponible actualmente para clústeres de EMR con varios nodos maestros ni para clústeres de EMR integrados con AWS Lake Formation.
- Para habilitar el acceso con un clic a las interfaces de usuario de las aplicaciones persistentes, debe tener permiso para acceder a Amazon EMR `DescribePersistentAppUI` y a `GetPersistentAppUIPresignedURL` las acciones de Amazon EMR. `CreatePersistentAppUI` Si deniega a un director de IAM el permiso para realizar estas acciones, el cambio de permiso tardará aproximadamente cinco minutos en propagarse.
- Si un clúster está habilitado para un rol en tiempo de ejecución, al acceder al Spark History Server desde la interfaz de usuario de la aplicación persistente, el usuario solo podrá acceder a un trabajo de Spark si el trabajo de Spark lo envía un rol en tiempo de ejecución.
- Si un clúster está habilitado para un rol en tiempo de ejecución, cada usuario solo podrá acceder a una solicitud enviada por la misma identidad de usuario y rol de tiempo de ejecución.
- La `AccessAllEventLogs` acción de Amazon EMR es necesaria para ver todas las aplicaciones en las interfaces de usuario de aplicaciones persistentes para los clústeres con funciones de tiempo de ejecución habilitadas.
- Si vuelve a configurar las aplicaciones en un clúster en ejecución, el historial de aplicaciones no estará disponible a través de la interfaz de usuario de la aplicación.
- Para cada uno de ellos Cuenta de AWS, el límite predeterminado de aplicaciones activas UIs es de 200.
- A continuación Regiones de AWS, puede acceder a la aplicación UIs desde la consola con Amazon EMR 6.14.0 y versiones posteriores:
  - Asia-Pacífico (Yakarta) (ap-southeast-3)
  - Europa (España) (eu-south-2)
  - Asia Pacífico (Melbourne) (ap-southeast-4)
  - Israel (Tel Aviv) (il-central-1)
  - Medio Oriente (EAU) (me-central-1)
- A continuación Regiones de AWS, puede acceder a la aplicación UIs desde la consola con Amazon EMR 5.25.0 y versiones posteriores:
  - Este de EE. UU. (Norte de Virginia) (us-east-1)
  - Oeste de EE. UU. (Oregón) (us-west-2)
  - Asia Pacífico (Bombay) (ap-south-1)
  - Asia-Pacífico (Seúl) (ap-northeast-2)
  - Asia-Pacífico (Singapur) (ap-southeast-1)

- Asia-Pacífico (Sídney) (ap-southeast-2)
- Asia-Pacífico (Tokio) (ap-northeast-1)
- Canadá (centro) (ca-central-1)
- América del Sur (São Paulo) (sa-east-1)
- Europa (Fráncfort) (eu-central-1)
- Europa (Irlanda) (eu-west-1)
- Europa (Londres) (eu-west-2)
- UE (París) (eu-west-3)
- Europa (Estocolmo) (eu-north-1)
- China (Pekín) (cn-north-1)
- China (Ningxia) (cn-northwest-1)

Visualización del historial de una aplicación de alto nivel en Amazon EMR

> **Note**
>
> Recomendamos que utilice la interfaz de aplicaciones persistentes para mejorar la experiencia de usuario y retener el historial de las aplicaciones durante un máximo de 30 días. El historial de aplicaciones de alto nivel que se describe en esta página no está disponible en la nueva consola de Amazon EMR (https://console.aws.amazon.com/emr). Para obtener más información, consulte Visualización de interfaces de usuario de aplicaciones persistentes en Amazon EMR.

Con las versiones de 5.8.0 a 5.36.0 y de 6.x a 6.8.0 de Amazon EMR, puede ver el historial de aplicaciones de alto nivel desde la pestaña Interfaces de usuario de aplicaciones de la consola antigua de Amazon EMR. La interfaz de usuario de aplicaciones de Amazon EMR conserva el resumen del historial de la aplicación durante 7 días después de que se haya completado la solicitud.

Consideraciones y limitaciones

Tenga en cuenta las siguientes limitaciones cuando utilice la pestaña Interfaces de usuario de aplicaciones en la antigua consola de Amazon EMR.

- Solo puede acceder a la característica del historial de aplicaciones de alto nivel cuando utiliza las versiones de 5.8.0 a 5.36.0 y de 6.x a 6.8.0 de Amazon EMR. A partir del 23 de enero de 2023,

Amazon EMR retirará el historial de aplicaciones de alto nivel para todas las versiones. Si utiliza la versión 5.25.0 de Amazon EMR o una superior, le recomendamos que utilice en su lugar la interfaz de usuario de aplicaciones persistentes.

- La característica del historial de aplicaciones de alto nivel no es compatible con las aplicaciones de Spark Streaming.
- El acceso con un solo clic a las interfaces de usuario de aplicaciones persistentes no está disponible actualmente para los clústeres de Amazon EMR con varios nodos maestros ni para los clústeres de Amazon EMR integrados con AWS Lake Formation.

Ejemplo: ver el historial de una aplicación de alto nivel

En la siguiente secuencia se muestra cómo desplazarse por los detalles de los trabajos en las aplicaciones de Spark o YARN utilizando la opción Interfaces de usuario de aplicaciones de la página de detalles del clúster de la consola antigua.

Para ver los detalles del clúster, seleccione Nombre en la lista Clústeres. Para ver información sobre los registros de contenedor de YARN, debe habilitar el registro para el clúster. Para más información, consulte Configurar el registro y la depuración de un clúster. Para el historial de aplicaciones de Spark, la información proporcionada en la tabla de resumen es solo un subconjunto de la información disponible a través de la interfaz de usuario del servidor del historial de Spark.

En la pestaña Interfaces de usuario de aplicaciones de Historial de aplicaciones de alto nivel, puede expandir una fila para mostrar el resumen del diagnóstico de una aplicación de Spark o seleccionar un enlace con el identificador de la aplicación para ver los detalles de otra aplicación.

Al seleccionar un enlace con un ID de aplicación, la interfaz de usuario cambia para mostrar los detalles de la aplicación YARN de esa aplicación. En la pestaña Trabajos de los detalles de la aplicación YARN, puede seleccionar el enlace Descripción de un trabajo para mostrar los detalles de ese trabajo.

En la página de detalles del trabajo, puede expandir la información sobre etapas específicas del trabajo y, a continuación, seleccione el enlace Descripción para ver los detalles de la etapa.

En la página de detalles de la etapa, puede ver las métricas clave de las tareas y los ejecutores de la etapa. También puede ver los registros de tareas y ejecutores mediante los enlaces Ver registros.

Visualización de los archivos de registro de Amazon EMR

Amazon EMR y Hadoop producen archivos de registro que notifican el estado en el clúster. De forma predeterminada, se escriben en el nodo principal del directorio `/mnt/var/log/`. En función de cómo configure el clúster al lanzarlo, estos registros también pueden archivarse en Amazon S3 y pueden ser visibles a través de la herramienta de depuración gráfica.

Existen muchos tipos de registros escritos en el nodo principal. Amazon EMR escribe registros de estado de instancias, pasos y acciones de arranque. Apache Hadoop escribe registros para notificar el procesamiento de los trabajos, de las tareas y de los intentos de tareas. Hadoop también guarda registros de sus daemons. Para obtener más información sobre los registros escritos por Hadoop, visite http://hadoop.apache. org/docs/stable/hadoop-project-dist/hadoop-common/ClusterSetup.html.

Ver archivos de registro en el nodo principal

En la siguiente tabla se enumeran algunos de los archivos de registro que encontrará en el nodo principal.

Ubicación	Descripción
/-acciones emr/instance-controller/log/bootstrap	Los registros escritos durante el procesamiento de las acciones de arranque.
/-impulsor mnt/var/log/hadoop de estados	Los registros escritos por el proceso state pusher de Hadoop.
/emr/instance-controller/log	Registros de controlador de instancia.
/emr/instance-state	Registros de estado de instancia. Contienen información sobre la CPU, el estado de la memoria y los subprocesos del recolector de elementos no utilizados del nodo.
/emr/service-nanny	Los registros escritos por el proceso nanny de servicio.
/mnt/var/log/*application*	Registros específicos de una aplicación como, por ejemplo, Hadoop, Spark o Hive.

Ubicación	Descripción
/mnt/var/log/hadoop/steps/*N*	Registros de paso que contienen informaci ón sobre el procesamiento del paso. El valor de *N* indica el StepID asignado por Amazon EMR. Por ejemplo, un clúster tiene dos pasos: `s-1234ABCDEFGH` y `s-5678IJKLMNOP` . El primer paso está ubicado en `/mnt/var/log/hadoop/steps/s-1234ABCDEFGH/` y el segundo paso en `/mnt/var/log/hadoop/steps/s-5678IJKLMNOP/` . Los registros de pasos escritos por Amazon EMR son los siguientes. • controlador: información sobre el procesami ento del paso. Si se produce un error en el paso durante la carga, puede encontrar el registro de seguimiento de la pila en este registro. • syslog: describe la ejecución de los trabajos de Hadoop en el paso. • stderr: el canal de error estándar de Hadoop mientras procesa el paso. • stdout: el canal de salida estándar de Hadoop mientras procesa el paso.

Para ver los archivos de registro en el nodo principal con la AWS CLI.

1. Utilice SSH para conectarse al nodo principal como se describe en Conexión al nodo principal del clúster de Amazon EMR mediante SSH.
2. Vaya al directorio que contiene la información del archivo de registro que desea ver. La tabla anterior ofrece una lista de los tipos de archivos de registro que están disponibles y donde los encontrará. El siguiente ejemplo muestra el comando para acceder al registro de paso con un ID, `s-1234ABCDEFGH`.

```
cd /mnt/var/log/hadoop/steps/s-1234ABCDEFGH/
```

3. Utilice el visor de archivos que desee para ver el archivo de log. En el siguiente ejemplo se usa el comando `less` de Linux para ver el archivo de log `controller`.

```
less controller
```

Ver los archivos de registro archivados en Amazon S3

De forma predeterminada, los clústeres de Amazon EMR que se lanzan utilizando la consola archivan automáticamente los archivos de registro en Amazon S3. Puede especificar su propia ruta de registro o bien puede permitir que la consola genere automáticamente una ruta de registro por usted. En los clústeres que se lanzan con la CLI o la API, debe configurar el archivado de registros de Amazon S3 manualmente.

Cuando Amazon EMR está configurado para archivar archivos de registro en Amazon S3, almacena los archivos en la ubicación S3 que especificó, en la carpeta/*cluster-id*/, donde *cluster-id* se encuentra el ID del clúster.

En la siguiente tabla se enumeran algunos de los archivos de registro que encontrará en Amazon S3.

Ubicación	Descripción
//node/ *cluster-id*	Los registros de nodo, incluida la acción de arranque, el estado de la instancia y los registros de aplicación para el nodo. Los registros de cada nodo se almacenan en una carpeta etiquetada con el identificador de la EC2 instancia de ese nodo.
/*cluster-id instance-id* / node//*application*	Los registros creados por cada aplicació n o daemon asociado con una aplicació n. Por ejemplo, el registro del servidor de Hive se encuentra en *cluster-i d* /node/*instance-id* /hive/hive-server.log .

Ubicación	Descripción
//pasos//*cluster-id step-id*	Registros de paso que contienen información sobre el procesamiento del paso. El valor de *step-id* indica el ID de paso asignado por Amazon EMR. Por ejemplo, un clúster tiene dos pasos: `s-1234ABCDEFGH` y `s-5678IJKLMNOP`. El primer paso está ubicado en `/mnt/var/log/hadoop/steps/s-1234ABCDEFGH/` y el segundo paso en `/mnt/var/log/hadoop/steps/s-5678IJKLMNOP/`. Los registros de pasos escritos por Amazon EMR son los siguientes. • controlador: información sobre el procesamiento del paso. Si se produce un error en el paso durante la carga, puede encontrar el registro de seguimiento de la pila en este registro. • syslog: describe la ejecución de los trabajos de Hadoop en el paso. • stderr: el canal de error estándar de Hadoop mientras procesa el paso. • stdout: el canal de salida estándar de Hadoop mientras procesa el paso.
//contenedores *cluster-id*	Registros de contenedor de aplicaciones. Los registros para cada aplicación YARN se almacenan en estas ubicaciones.
//hadoop-mapreduce/ *cluster-id*	Los registros que contienen información sobre los detalles de configuración y el historial de tareas de las tareas. MapReduce

Para ver los archivos de registro archivados en Amazon S3 con la consola de Amazon S3

1. Inicie sesión en la consola de Amazon S3 AWS Management Console y ábrala en https://console.aws.amazon.com/s3/.
2. Abra el bucket de S3 que especificó al configurar el clúster para archivar los archivos de registro en Amazon S3.
3. Acceda al archivo de registro que contiene la información que se va a mostrar. La tabla anterior ofrece una lista de los tipos de archivos de registro que están disponibles y donde los encontrará.
4. Descargue el objeto del archivo de registro para verlo. Para obtener instrucciones, consulte Descarga de un objeto.

Ver instancias de clúster en Amazon EC2

Para ayudarte a gestionar tus recursos, Amazon te EC2 permite asignar metadatos a los recursos en forma de etiquetas. Cada EC2 etiqueta de Amazon consta de una clave y un valor. Las etiquetas te permiten clasificar tus EC2 recursos de Amazon de diferentes maneras: por ejemplo, por propósito, propietario o entorno.

Puede buscar y filtrar los recursos en función de las etiquetas. Las etiquetas que asignas a los recursos a través de tu AWS cuenta solo están disponibles para ti. Otras cuentas que compartan el mismo recurso no pueden ver sus etiquetas.

Amazon EMR etiqueta automáticamente cada EC2 instancia que lanza con pares clave-valor. Las claves identifican el clúster y el grupo de instancias al que pertenece la instancia. Esto facilita el filtrado de EC2 las instancias para mostrar, por ejemplo, solo las instancias que pertenecen a un clúster en particular o para mostrar todas las instancias que se están ejecutando actualmente en el grupo de instancias de la tarea. Esto resulta especialmente útil si ejecutas varios clústeres de forma simultánea o si administras un gran número de EC2 instancias.

Estos son los pares clave-valor predefinidos que asigna Amazon EMR:

Clave	Valor	Definición de los valores
aws:elastic mapreduce:job-flow-id	*job-flow-identifier*	El ID del clúster para el que se aprovisiona la instancia. Aparece en el formato `j-XXXXXXXXXXXXX` y puede tener una longitud de hasta 256 caracteres.

Clave	Valor	Definición de los valores
aws:elasticmapreduce:instance-group-role	*group-role*	El tipo de grupo de instancias, introducido como uno de los siguientes valores: `master`, `core` o `task`.

Puede ver y filtrar por las etiquetas que agrega Amazon EMR. Para obtener más información, consulta Uso de etiquetas en la Guía del EC2 usuario de Amazon. Dado que las etiquetas definidas por Amazon EMR son etiquetas del sistema y no se pueden editar ni eliminar, las secciones sobre mostrar y filtrar las etiquetas son las más importantes.

> **Note**
>
> Amazon EMR añade etiquetas a la EC2 instancia cuando su estado se actualiza a En ejecución. Si la latencia se produce entre el momento en que se aprovisiona la EC2 instancia y el momento en que su estado se establece en En ejecución, las etiquetas que establezca Amazon EMR aparecerán una vez que se inicie la instancia. Si no ve las etiquetas, espere algunos minutos y actualice la vista.

CloudWatch eventos y métricas de Amazon EMR

Utilice eventos y métricas para realizar un seguimiento de la actividad y el estado de un clúster de Amazon EMR. Los eventos son útiles para monitorizar la incidencia específica dentro de un clúster; por ejemplo, cuando un clúster cambia de estado desde el inicio a la ejecución. Las métricas son útiles para monitorizar un valor específico; por ejemplo, el porcentaje de espacio en disco disponible que utiliza HDFS dentro de un clúster.

Para obtener más información sobre CloudWatch los eventos, consulte la Guía del usuario de Amazon CloudWatch Events. Para obtener más información sobre CloudWatch las métricas, consulta Uso de CloudWatch métricas de Amazon y Creación de CloudWatch alarmas de Amazon en la Guía del CloudWatch usuario de Amazon.

Temas

- Supervisión de las métricas de Amazon EMR con CloudWatch
- Supervisión de eventos de Amazon EMR con CloudWatch

- Respuesta a CloudWatch eventos de Amazon EMR

Supervisión de las métricas de Amazon EMR con CloudWatch

Las métricas se actualizan cada cinco minutos y se recopilan y envían automáticamente CloudWatch para cada clúster de Amazon EMR. Este intervalo no se puede configurar. No se cobran cargos por las métricas de Amazon EMR informadas en. CloudWatch Estas métricas de puntos de datos de cinco minutos se archivan durante 63 días, tras lo cual se descartan los datos.

¿Cómo utilizo las métricas de Amazon EMR?

En la siguiente tabla, se muestran los usos más comunes de las métricas notificadas por Amazon EMR. Se trata de sugerencias que puede usar como punto de partida y no de una lista completa. Para ver una lista completa de las métricas notificadas por Amazon EMR, consulte Métricas reportadas por Amazon EMR en CloudWatch.

¿Cómo?	Métricas relevantes
Realizar un seguimiento del progreso de mi clúster	Consulte las métricas `RunningMapTasks` , `RemainingMapTasks` , `RunningReduceTasks` y `RemainingReduceTasks` .
Detectar clústeres que están inactivos	La métrica `IsIdle` realiza un seguimiento de si el clúster está disponible, pero actualmente no está ejecutando ninguna tarea. Puede configurar una alarma para que se active cuando el clúster haya estado inactivo durante un periodo de tiempo determinado, como, por ejemplo, treinta minutos.
Detectar si un nodo se queda sin espacio de almacenamiento	La métrica `MRUnhealthyNodes` registra cuándo uno o más nodos principales o de tarea se quedan sin almacenamiento en disco local y pasan a un estado `UNHEALTHY` de YARN. Por ejemplo, los nodos principales o de tarea se están quedando sin espacio en disco y no podrán ejecutar tareas.

¿Cómo?	Métricas relevantes
Detectar si un clúster se queda sin espacio de almacenamiento	La métrica `HDFSUtilization` monitorea la capacidad de HDFS combinada del clúster y puede requerir el redimensionamiento del clúster para agregar más nodos principales. Por ejemplo, el uso de HDFS es elevado, lo que podría afectar a los trabajos y al estado del clúster.
Detectar cuándo un clúster se está ejecutando a una capacidad reducida	La métrica `MRLostNodes` registra cuándo uno o más nodos principales o de tarea no pueden comunicarse con el nodo maestro. Por ejemplo, el nodo maestro no puede acceder al nodo principal o de tarea.

Para obtener más información, consulte El clúster de Amazon EMR finaliza con NO_SLAVE_LEFT y los nodos básicos con FAILED_BY_MASTER y AWSSupport-AnalyzeEMRLogs.

Accede a CloudWatch las métricas de Amazon EMR

Puede ver las métricas de las que informa Amazon EMR CloudWatch mediante la consola Amazon EMR o la consola. CloudWatch También puede recuperar métricas mediante el comando CloudWatch CLI `mon-get-stats` o la CloudWatch `GetMetricStatistics` API. Para obtener más información sobre cómo ver o recuperar las métricas de Amazon EMR CloudWatch mediante Amazon, consulte la Guía del usuario de CloudWatch Amazon.

Console

Para ver métricas con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija el clúster cuyas métricas desee ver. Se abrirá la página de detalles del clúster.
3. Seleccione la pestaña Monitorización en la página de detalles del clúster. Seleccione cualquiera de las opciones denominadas Estado del clúster, Estado del nodo o Entradas y salidas para cargar los informes acerca del progreso y el estado del clúster.

4. Después de seleccionar una métrica para verla, puede modificar el tamaño de cada gráfico. Para filtrar el periodo de tiempo del gráfico, seleccione una opción preconfigurada o seleccione Personalizado.

Métricas reportadas por Amazon EMR en CloudWatch

En las siguientes tablas se enumeran las métricas que Amazon EMR informa en la consola y las envía. CloudWatch

Métricas de Amazon EMR

Amazon EMR envía datos de varias métricas a. CloudWatch Todos los clústeres de Amazon EMR envían automáticamente métricas en intervalos de cinco minutos. Las métricas se archivan durante dos semanas; después de ese periodo, los datos se descartan.

El espacio de nombres de `AWS/ElasticMapReduce` incluye las siguientes métricas.

> **Note**
>
> Amazon EMR extrae métricas de un clúster. Si un clúster deja de estar disponible, no se registra ninguna métrica hasta que el clúster vuelve a estar disponible.

Están disponibles las siguientes métricas para los clústeres que ejecutan las versiones 2.x de Hadoop.

Métrica	Descripción
Estado del clúster	
IsIdle	Indica que un clúster ya no está funcionando, pero sigue activo y acumulando cargos. Se establece en 1 si no se ejecuta ninguna tarea ni ningún trabajo; en caso contrario , se establece en 0. Este valor se comprueba a intervalo s de cinco minutos, y un valor de 1 indica que el clúster estaba inactivo cuando se comprobó, no que estuvo inactivo durante los cinco minutos. Para evitar falsos positivos, debe activar una alarma cuando este valor sea 1 durante más de una comprobación consecutiva de cinco minutos. Por

Métrica	Descripción
	ejemplo, puede activar una alarma cuando este valor sea 1 durante treinta minutos o más. Caso de uso: monitorizar el rendimiento del clúster Unidades: booleano
ContainerAllocated	El número de contenedores de recursos asignados por. ResourceManager Caso de uso: monitorizar el progreso del clúster Unidades: recuento
ContainerReserved	El número de contenedores reservados. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
ContainerPending	El número de contenedores en la cola que aún no se han asignado. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
ContainerPendingRatio	La relación entre los contenedores pendientes y los contenedores asignados (ContainerPendingRatio = ContainerPending / ContainerAllocated). Si Container Allocated = 0, entonces ContainerPendingRatio =Containe rPending. El valor de ContainerPendingRatio representa un número, no un porcentaje. Este valor es útil para escalar recursos del clúster en función del comportamiento de asignación de contenedores. Unidades: recuento

Métrica	Descripción
AppsCompleted	El número de aplicaciones enviadas a YARN que se han completado. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
AppsFailed	El número de aplicaciones enviadas a YARN que no se han podido completar. Caso de uso: monitorizar el progreso del clúster, monitorizar el estado del clúster Unidades: recuento
AppsKilled	El número de aplicaciones enviadas a YARN que se han cancelado. Caso de uso: monitorizar el progreso del clúster, monitorizar el estado del clúster Unidades: recuento
AppsPending	El número de aplicaciones enviadas a YARN que están en estado pendiente. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
AppsRunning	El número de aplicaciones enviadas a YARN que se están ejecutando. Caso de uso: monitorizar el progreso del clúster Unidades: recuento

Métrica	Descripción
AppsSubmitted	El número de aplicaciones enviadas a YARN. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
Estado del nodo	
CoreNodesRunning	El número de nodos secundarios en funcionamiento. Los puntos de datos de esta métrica solo se registran cuando existe un grupo de instancias correspondiente. Caso de uso: monitorizar el estado del clúster Unidades: recuento
CoreNodesPending	El número de nodos secundarios en espera de ser asignados. Puede que no todos los nodos secundarios solicitados estén disponibles inmediatamente; esta métrica registra las solicitudes pendientes. Los puntos de datos de esta métrica solo se registran cuando existe un grupo de instancias correspondiente. Caso de uso: monitorizar el estado del clúster Unidades: recuento
LiveDataNodes	El porcentaje de nodos de datos que reciben trabajo de Hadoop. Caso de uso: monitorizar el estado del clúster Unidades: porcentaje

Métrica	Descripción
MRTotalNodos	El número de nodos actualmente disponibles para los MapReduce trabajos. Equivalente a la métrica YARN `mapred.resourcemanager.TotalNodes` . Caso de uso: monitorizar el progreso del clúster Unidades: recuento Nota: MRTotal Los nodos solo cuentan los nodos actualmen te activos en el sistema. YARN elimina automáticamente los nodos terminados de este recuento y deja de rastrearl os, por lo que no se consideran en la métrica de MRTotal nodos.
MRActiveNodos	El número de nodos que actualmente ejecutan MapReduce tareas o trabajos. Equivalente a la métrica YARN `mapred.resourcemanager.NoOfActiveNodes` . Caso de uso: monitorizar el progreso del clúster Unidades: recuento
MRLostNodos	El número de nodos asignados a los MapReduce que se ha marcado como PERDIDO. Equivalente a la métrica YARN `mapred.resourcemanager.NoOfLostNodes` . Caso de uso: monitorizar el estado del clúster, monitorizar el progreso del clúster Unidades: recuento

Métrica	Descripción
MRUnhealthyNodos	El número de nodos disponibles para los MapReduce trabajos marcados como insalubres. Equivalente a la métrica YARN `mapred.resourcemanager.NoOf UnhealthyNodes` . Caso de uso: monitorizar el progreso del clúster Unidades: recuento
MRDecommissionedNodos	El número de nodos asignados a MapReduce las aplicacio nes que se han marcado como desactivadas. Equivalente a la métrica YARN `mapred.resourcemanager.NoOf DecommissionedNodes` . Caso de uso: monitorizar el estado del clúster, monitorizar el progreso del clúster Unidades: recuento
MRRebootedNodos	El número de nodos disponibles MapReduce que se han reiniciado y marcado en estado REINICIADO. Equivalente a la métrica YARN `mapred.resourcemanager.NoOf RebootedNodes` . Caso de uso: monitorizar el estado del clúster, monitorizar el progreso del clúster Unidades: recuento
MultiMasterInstanceGroupNod esRunning	El número de nodos principales en ejecución. Caso de uso: monitorizar errores y sustituciones del nodo principal Unidades: recuento

Métrica	Descripción
MultiMasterInstanceGroupNodesRunningPercentage	El porcentaje de nodos principales que se están ejecutando o por encima del recuento de instancias del nodo principal solicitadas. Caso de uso: monitorizar errores y sustituciones del nodo principal Unidades: porcentaje
MultiMasterInstanceGroupNodesRequested	El número de nodos principales solicitados. Caso de uso: monitorizar errores y sustituciones del nodo principal Unidades: recuento
E/S	
S3 BytesWritten	Número de bytes escritos en Amazon S3. Esta métrica solo agrega MapReduce trabajos y no se aplica a otras cargas de trabajo en Amazon EMR. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento
S3 BytesRead	Número de bytes leídos desde Amazon S3. Esta métrica solo agrega MapReduce trabajos y no se aplica a otras cargas de trabajo en Amazon EMR. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento

Métrica	Descripción
HDFSUtilization	El porcentaje de almacenamiento HDFS usado actualmente. Caso de uso: analizar el rendimiento del clúster Unidades: porcentaje
HDFSBytesLea	El número de bytes leídos de HDFS. Esta métrica solo agrega MapReduce trabajos y no se aplica a otras cargas de trabajo en Amazon EMR. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento
HDFSBytesEscrito	El número de bytes escritos en HDFS. Esta métrica solo agrega MapReduce trabajos y no se aplica a otras cargas de trabajo en Amazon EMR. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento
MissingBlocks	El número de bloques en los que HDFS no tiene réplicas. Pueden tratarse de bloques dañados. Caso de uso: monitorizar el estado del clúster Unidades: recuento
CorruptBlocks	El número de bloques que HDFS registra como dañados. Caso de uso: monitorizar el estado del clúster Unidades: recuento

Métrica	Descripción
TotalLoad	El número total de transferencias de datos simultáneas, Caso de uso: monitorizar el estado del clúster Unidades: recuento
MemoryTotalMB	La cantidad de memoria total del clúster. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
MemoryReservedMB	La cantidad de memoria reservada. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
MemoryAvailableMB	La cantidad de memoria disponible para asignar. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
YARNMemoryAvailablePercentage	El porcentaje de memoria restante disponible para YARN (YARNMemoryAvailablePercentage = MemoryAvailable MB/ MemoryTotalMB). Este valor es útil para escalar recursos del clúster en función del uso de memoria de YARN. Unidades: porcentaje
MemoryAllocatedMB	La cantidad de memoria asignada al clúster. Caso de uso: monitorizar el progreso del clúster Unidades: recuento

Métrica	Descripción
PendingDeletionBlocks	El número de bloques marcados para eliminación. Caso de uso: monitorizar el progreso del clúster, monitorizar el estado del clúster Unidades: recuento
UnderReplicatedBlocks	El número de bloques que necesitan replicarse una o varias veces. Caso de uso: monitorizar el progreso del clúster, monitorizar el estado del clúster Unidades: recuento
DfsPendingReplicationBlocks	El estado de replicación del bloque: bloques que se están replicando, antigüedad de las solicitudes de replicación y solicitudes replicadas correctamente. Caso de uso: monitorizar el progreso del clúster, monitorizar el estado del clúster Unidades: recuento
CapacityRemainingGB	La cantidad de capacidad de disco HDFS restante. Caso de uso: monitorizar el progreso del clúster, monitorizar el estado del clúster Unidades: recuento

A continuación se indican las métricas de Hadoop 1:

Métrica	Descripción
Estado del clúster	

Métrica	Descripción
IsIdle	Indica que un clúster ya no está funcionando, pero sigue activo y acumulando cargos. Se establece en 1 si no se ejecuta ninguna tarea ni ningún trabajo; en caso contrario , se establece en 0. Este valor se comprueba a intervalo s de cinco minutos, y un valor de 1 indica que el clúster estaba inactivo cuando se comprobó, no que estuvo inactivo durante los cinco minutos. Para evitar falsos positivos, debe activar una alarma cuando este valor sea 1 durante más de una comprobación consecutiva de cinco minutos. Por ejemplo, puede activar una alarma cuando este valor sea 1 durante treinta minutos o más. Caso de uso: monitorizar el rendimiento del clúster Unidades: booleano
JobsRunning	El número de trabajos del clúster que se encuentran actualmente en ejecución. Caso de uso: monitorizar el estado del clúster Unidades: recuento
JobsFailed	El número de trabajos del clúster que han producido un error. Caso de uso: monitorizar el estado del clúster Unidades: recuento
Map/Reduce	

Métrica	Descripción
MapTasksRunning	Número de tareas de asignación en ejecución de cada trabajo. Si tiene un programador instalado y varios trabajos en ejecución, se generan varios gráficos. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
MapTasksRemaining	Número de tareas de asignación pendientes de cada trabajo. Si tiene un programador instalado y varios trabajos en ejecución, se generan varios gráficos. Una tarea de asignación pendiente es aquella que no tiene ninguno de los siguientes estados: Running, Killed o Completed. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
MapSlotsOpen	La capacidad de la tarea de asignación no utilizada. Se calcula como el número máximo de tareas de asignación de un clúster determinado menos el número total de tareas de asignación que se están ejecutando actualmente en el clúster. Caso de uso: analizar el rendimiento del clúster Unidades: recuento
RemainingMapTasksPerSlot	La proporción entre el total de tareas de asignación pendientes y el total de slots de asignación disponibles en el clúster. Caso de uso: analizar el rendimiento del clúster Unidades: proporción

Métrica	Descripción
ReduceTasksRunning	Número de tareas de reducción en ejecución de cada trabajo. Si tiene un programador instalado y varios trabajos en ejecución, se generan varios gráficos. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
ReduceTasksRemaining	Número de tareas de reducción pendientes de cada trabajo. Si tiene un programador instalado y varios trabajos en ejecución, se generan varios gráficos. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
ReduceSlotsOpen	La capacidad de la tarea de reducción no utilizada. Se calcula como la capacidad máxima de la tarea de reducción de un clúster determinado menos el número de tareas de reducción que se están ejecutando actualmente en el clúster. Caso de uso: analizar el rendimiento del clúster Unidades: recuento
Estado del nodo	
CoreNodesRunning	El número de nodos secundarios en funcionamiento. Los puntos de datos de esta métrica solo se registran cuando existe un grupo de instancias correspondiente. Caso de uso: monitorizar el estado del clúster Unidades: recuento

Métrica	Descripción
CoreNodesPending	El número de nodos secundarios en espera de ser asignados. Puede que no todos los nodos secundarios solicitados estén disponibles inmediatamente; esta métrica registra las solicitudes pendientes. Los puntos de datos de esta métrica solo se registran cuando existe un grupo de instancias correspondiente. Caso de uso: monitorizar el estado del clúster Unidades: recuento
LiveDataNodes	El porcentaje de nodos de datos que reciben trabajo de Hadoop. Caso de uso: monitorizar el estado del clúster Unidades: porcentaje
TaskNodesRunning	El número de nodos de tareas en funcionamiento. Los puntos de datos de esta métrica solo se registran cuando existe un grupo de instancias correspondiente. Caso de uso: monitorizar el estado del clúster Unidades: recuento
TaskNodesPending	El número de nodos de tareas en espera de ser asignados . Puede que no todos los nodos de tareas solicitados estén disponibles inmediatamente; esta métrica registra las solicitudes pendientes. Los puntos de datos de esta métrica solo se registran cuando existe un grupo de instancias correspondiente. Caso de uso: monitorizar el estado del clúster Unidades: recuento

Métrica	Descripción
LiveTaskTrackers	El porcentaje de rastreadores de tareas que están operativos. Caso de uso: monitorizar el estado del clúster Unidades: porcentaje
E/S	
S3 BytesWritten	Número de bytes escritos en Amazon S3. Esta métrica solo agrega MapReduce trabajos y no se aplica a otras cargas de trabajo en Amazon EMR. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento
S3 BytesRead	Número de bytes leídos desde Amazon S3. Esta métrica solo agrega MapReduce trabajos y no se aplica a otras cargas de trabajo en Amazon EMR. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento
HDFSUtilization	El porcentaje de almacenamiento HDFS usado actualmente. Caso de uso: analizar el rendimiento del clúster Unidades: porcentaje
HDFSBytesLea	El número de bytes leídos de HDFS. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento

Métrica	Descripción
HDFSBytesEscrito	El número de bytes escritos en HDFS. Caso de uso: analizar el rendimiento del clúster, monitorizar el progreso del clúster Unidades: recuento
MissingBlocks	El número de bloques en los que HDFS no tiene réplicas. Pueden tratarse de bloques dañados. Caso de uso: monitorizar el estado del clúster Unidades: recuento
TotalLoad	El número total actual de lectores y escritores registrado por todos los miembros DataNodes de un grupo. Caso de uso: Diagnostique en qué medida una dosis alta I/O podría estar contribuyendo a un desempeño deficient e en la ejecución del trabajo. Los nodos de trabajo que ejecutan el DataNode daemon también deben realizar tareas de mapeo y reducción. La persistencia de TotalLoad valores altos a lo largo del tiempo puede indicar que un nivel alto I/O podría ser un factor que contribuya a un rendimiento deficiente. Los picos ocasionales en este valor son normales y no necesariamente son indicativos de un problema. Unidades: recuento

Métricas de capacidad del clúster

Las siguientes métricas indican la capacidad actual o prevista de un clúster. Estas métricas solo están disponibles cuando el escalado administrado o la terminación automática están habilitados.

Para los clústeres compuestos por flotas de instancias, las métricas de capacidad del clúster se miden en `Units`. Para los clústeres compuestos por grupos de instancias, las métricas de capacidad

del clúster se miden en `Nodes` o en `VCPU` en función del tipo de unidad utilizado en la política de escalado administrado. Para obtener más información, consulte Uso del escalado administrado de EMR en la Guía de administración de Amazon EMR.

Métrica	Descripción
• `TotalUnitsRequested` • `TotalNodesRequested` • `TotalVCPURequested`	El número total objetivo de units/nodes/vCPUs un clúster, determinado por el escalado gestionado. Unidades: recuento
• `TotalUnitsRunning` • `TotalNodesRunning` • `TotalVCPURunning`	El número total actual de units/nodes/vCPUs disponibles en un clúster en ejecución. Cuando se solicita un cambio de tamaño del clúster, esta métrica se actualizará después de agregar o quitar las nuevas instancias del clúster. Unidades: recuento
• `CoreUnitsRequested` • `CoreNodesRequested` • `CoreVCPURequested`	El número objetivo de CORE units/nodes/vCPUs en un clúster, determinado por el escalado gestionado. Unidades: recuento
• `CoreUnitsRunning` • `CoreNodesRunning` • `CoreVCPURunning`	El número actual de CORE units/nodes/vCPUs que se ejecutan en un clúster. Unidades: recuento
• `TaskUnitsRequested`	El número objetivo de TAREAS units/nodes/vCPUs en un clúster, determinado por el escalado gestionado.

Métrica	Descripción
• `TaskNodesRequested` • `TaskVCPURequested`	Unidades: recuento
• `TaskUnitsRunning` • `TaskNodesRunning` • `TaskVCPURunning`	El número actual de TAREAS units/nodes/vCPUs que se están ejecutando en un clúster. Unidades: recuento

Amazon EMR emite las siguientes métricas con una granularidad de un minuto cuando se habilita la terminación automática mediante una política de terminación automática. Algunas métricas solo están disponibles para las versiones 6.4.0 y posteriores de Amazon EMR. Para obtener más información acerca de la terminación automática, consulte Uso de una política de finalización automática para la limpieza de clústeres de Amazon EMR..

Métrica	Descripción
`TotalNotebookKernels`	El número total de kernels de cuadernos en ejecución e inactivos del clúster. Esta métrica solo está disponible para las versiones 6.4.0 y posteriores de Amazon EMR.
`AutoTerminationIsClusterIdle`	Indica si el clúster está en uso. Un valor de 0 indica que el clúster está siendo utilizado activamente por uno de los siguientes componentes: • Una aplicación YARN

Métrica	Descripción
	• HDFS • Un cuaderno • Una interfaz de usuario integrada en el clúster, como el servidor del historial de Spark Un valor de 1 indica que el clúster está inactivo. Amazon EMR comprueba la inactividad continua de los clústeres (`AutoTerminationIsClusterIdle` = 1). Cuando el tiempo de inactividad de un clúster es igual al valor de `IdleTimeout` de su política de terminación automática, Amazon EMR termina el clúster.

Dimensiones para las métricas de Amazon EMR

Los datos de Amazon EMR se pueden filtrar mediante alguna de las dimensiones de la tabla siguiente.

Dimensión	Descripción
JobFlowId	El mismo que el ID del clúster, que es un identificador único de un clúster con el formato `j-XXXXXXXXXXXXX`. Puede encontrar este valor haciendo clic en el clúster en la consola de Amazon EMR.

Supervisión de eventos de Amazon EMR con CloudWatch

Amazon EMR realiza un seguimiento de eventos y mantiene información acerca de ellos durante un máximo de siete días en la consola de Amazon EMR. Amazon EMR registra los eventos cuando se

produce un cambio en el estado de los clústeres, los grupos de instancias, las flotas de instancias, las políticas de escalado automático o los pasos. Los eventos capturan la fecha y la hora en que se produjo el evento, los detalles sobre los elementos afectados y otros puntos de datos críticos.

En la siguiente tabla se muestran eventos de Amazon EMR, junto con el estado o cambio de estado que indica el evento, la gravedad del evento, el tipo del evento, el código del evento y los mensajes de eventos. Amazon EMR representa los eventos como objetos JSON y los envía automáticamente a una secuencia de eventos. El objeto JSON es importante a la hora de configurar reglas para el procesamiento de CloudWatch eventos mediante Events, ya que las reglas buscan hacer coincidir los patrones del objeto JSON. Para obtener más información, consulte Eventos y patrones de eventos y Eventos de Amazon EMR en la Guía del usuario de Amazon CloudWatch Events.

> **Note**
>
> EMR emite eventos periódicamente con el EC2 aprovisionamiento del código del evento: capacidad de instancia insuficiente. Estos eventos se producen cuando su clúster de Amazon EMR detecta un error de capacidad insuficiente por parte de Amazon EMR para su flota de instancias o grupo de instancias durante la creación del clúster o la operación de cambio de tamaño. Es posible que un evento no incluya todos los tipos de instancias AZs que usted ha proporcionado, ya que EMR solo incluye los tipos de instancia e AZs intentó aprovisionar capacidad desde la última vez que se emitió el evento de capacidad insuficiente. Para obtener información sobre cómo responder a estos eventos, consulte Responder a eventos de capacidad de instancias insuficiente en el clúster de Amazon EMR.

Eventos de inicio de clústeres

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisio namiento de la flota de instancia s de EMR	EC2 aprovisio namiento: capacidad de instancia insuficiente	No podemos crear su clúster de Amazon EMR para Instance Fleet Amazon `InstanceFleetID` no EC2 tiene

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				suficiente capacidad de puntos `ClusterId (ClusterName)` para el tipo de instancia ni suficiente capacidad bajo demanda para el tipo `[Instancetype1, Instancetype2]` de instancia `[Instancetype3, Instancetype4]` en la zona de disponibilidad. `[AvailabilityZone1, AvaliabilityZone2]` Consulte aquí la documentación para obtener más información sobre cómo

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				responder a este evento.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisionamiento del grupo de instancias de EMR	EC2 aprovisionamiento: capacidad de instancia insuficiente	No podemos crear su clúster de Amazon EMR `ClusterId (ClusterName)` para el grupo de instancias. `InstanceGroupID` Amazon no EC2 tiene suficiente capacidad de puntos para el tipo de instancia ni suficiente capacidad bajo demanda para el tipo `[Instancetype1, Instancetype2]` de instancia `[Instancetype3, Instancetype4]` en la zona de disponibilidad. `[AvailabilityZone1, Avaliab`

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				`ilityZone 2]` Consulte aquí la documentación para obtener más informaci ón sobre cómo responder a este evento.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisio namiento de la flota de instancia s de EMR	EC2 aprovisio namiento: direcciones libres insuficie ntes en la subred	No podemos crear el clúster de Amazon EMR `ClusterId` `(ClusterN ame)` que ha solicitado para la flota de instancia s `InstanceF leetID` porque la subred especificada `[Subnet1, Subnet2]` no contiene suficientes direcciones IP privadas libres para atender su solicitud. Utilice la operación `DescribeS ubnets` para ver cuántas direccion es IP están disponibles (no se utilizan) en la subred. Para obtener información sobre cómo responder a este

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				evento, consulta Códigos de error de la EC2 API de Amazon

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisionamiento del grupo de instancias de EMR	EC2 aprovisionamiento: direcciones libres insuficientes en la subred	No podemos crear el clúster de Amazon EMR `ClusterId` `(ClusterName)` que solicitó para el grupo de instancias `InstanceGroupID` porque la subred especificada `[Subnet1, Subnet2]` no contiene suficientes direcciones IP privadas libres para atender su solicitud. Utilice la operación `DescribeSubnets` para ver cuántas direcciones IP están disponibles (no se utilizan) en la subred. Para obtener información sobre cómo

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				responder a este evento, consulta Códigos de error de la EC2 API de Amazon
`CREATING`	`WARN`	Aprovisionamiento de la flota de instancias de EMR	EC2 Aprovisionamiento: se ha superado el límite de vCPU	El aprovisionamiento del `InstanceFleetID` clúster de Amazon EMR `ClusterId` `(ClusterName)` se retrasa porque ha alcanzado el límite de vCPUs (unidades de procesamiento virtuales) asignadas a las instancias en ejecución de su. `account` `(accountId)` Para obtener más información, Códigos de error para la EC2 API de Amazon

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisio namiento del grupo de instancias de EMR	EC2 Aprovisio namiento: se ha superado el límite de vCPU	El aprovisio namiento del grupo de instancias `InstanceG roupID` en el clúster de Amazon EMR `ClusterId` se retrasa porque ha alcanzado el límite de v CPUs (unidades de procesami ento virtuales) asignadas a las instancias en ejecución de su cuenta. `(accountId)` Para obtener más información, Códigos de error para la EC2 API de Amazon

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisio namiento de la flota de instancia s de EMR	EC2 Aprovisio namiento: se ha superado el límite de número de instancias puntuales	El aprovisio namiento de la flota de instancia s `InstanceF leetID` en el clúster de Amazon EMR `ClusterID (ClusterN ame)` se retrasa porque ha alcanzado el límite del número de instancias de Spot que usted puede lanzar en su `account (accountId)` . Para obtener más informaci ón, consulta Códigos de error de la EC2 API de Amazon.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisionamiento del grupo de instancias de EMR	EC2 Aprovisionamiento: se ha superado el límite de recuento de instancias puntuales	El aprovisionamiento del grupo de instancias `InstanceGroupID` en el clúster de Amazon EMR `ClusterID` `(ClusterName)` se retrasa porque ha alcanzado el límite del número de instancias de Spot que puede lanzar en su `account` `(accountId)`. Para obtener más información, consulta Códigos de error de la EC2 API de Amazon.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisionamiento de la flota de instancias de EMR	EC2 Aprovisionamiento: se ha superado el límite de instancias	El aprovisionamiento de la flota de instancias `InstanceFleetID` en el clúster de Amazon EMR `ClusterId (ClusterName)` se retrasa porque ha alcanzado el límite del número de instancias que usted puede lanzar simultáneamente en su `account (accountID)`. Para obtener más información sobre los límites EC2 del servicio de Amazon, consulta Códigos de error para la EC2 API de Amazon.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisionamiento del grupo de instancias de EMR	EC2 Aprovisionamiento: se ha superado el límite de instancias	El aprovisionamiento del grupo de instancias `InstanceGroupID` en el clúster de Amazon EMR `ClusterId` `(ClusterName)` se retrasa porque ha alcanzado el límite de instancias que puede ejecutar simultáneamente en su `account` `(accountID)` . Para obtener más información sobre los límites EC2 del servicio de Amazon, consulta Códigos de error para la EC2 API de Amazon.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`CREATING`	`WARN`	Aprovisionamiento del grupo de instancias de EMR	ninguno	El clúster de Amazon EMR `ClusterId (ClusterName)` se creó a las `Time` y está listo para su uso. -o bien- El clúster de Amazon EMR `ClusterId (ClusterName)` terminó de ejecutar todos los pasos pendientes a las `Time`. Note Un clúster en estado `WAITING` todavía podría estar procesando trabajos.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`STARTING`	`INFO`	Cambio de estado del clúster de EMR	ninguno	El clúster de Amazon EMR `ClusterId (ClusterName)` se solicitó a las `Time` y se está creando.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`STARTING`	`INFO`	Cambio de estado del clúster de EMR	ninguno	**Note** Se aplica solo a los clústeres con la configuración de flotas de instancias y varias zonas de disponibilidad seleccionadas en Amazon EC2. El clúster de Amazon EMR `ClusterId (ClusterName)` se está creando en la zona (`AvailabilityZoneID`), que se ha elegido a partir

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				de las opciones de zona de disponibilidad especificadas.
`STARTING`	`INFO`	Cambio de estado del clúster de EMR	ninguno	El clúster de Amazon EMR `ClusterId (ClusterName)` comenzó a ejecutar los pasos a las `Time`.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`WAITING`	`INFO`	Cambio de estado del clúster de EMR	ninguno	El clúster de Amazon EMR `ClusterId (ClusterName)` se creó a las `Time` y está listo para su uso. -o bien- El clúster de Amazon EMR `ClusterId (ClusterName)` terminó de ejecutar todos los pasos pendientes a las `Time`. Note Un clúster en estado `WAITING` todavía podría estar procesando trabajos.

> **Note**
>
> Los eventos con código de evento se emiten `EC2 provisioning - Insufficient Instance Capacity` periódicamente cuando tu clúster de EMR detecta un error de capacidad insuficiente por parte de Amazon EC2 para tu flota de instancias o grupo de instancias durante la creación del clúster o la operación de cambio de tamaño. Para obtener información sobre cómo responder a estos eventos, consulte Responder a eventos de capacidad de instancias insuficiente en el clúster de Amazon EMR.

Eventos de terminación de clústeres

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`TERMINATED`	La gravedad depende del motivo por el que cambia el estado, tal y como se muestra a continuación: • **`CRITICAL`** si el clúster terminó con cualquiera de los siguientes motivos de cambio de estado: `INTERNAL_ERROR`, `VALIDATION_ERROR`, `INSTANCE_`	Cambio de estado del clúster de EMR	ninguno	El clúster de Amazon EMR `ClusterId (ClusterName)` terminó a las `Time` con el siguiente motivo: `StateChangeReason: Code` .

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
	`FAILURE` , `BOOTSTRAP_FAILURE` o `STEP_FAILURE` . • **`INFO`** si el clúster terminó con cualquiera de los siguientes motivos de cambio de estado: `USER_REQUEST` o `ALL_STEPS_COMPLETED` .			
`TERMINATED_WITH_ERRORS`	`CRITICAL`	Cambio de estado del clúster de EMR	ninguno	El clúster de Amazon EMR `ClusterId` `(ClusterName)` terminó con errores a las `Time` con el siguiente motivo: `StateChangeReason: Code` .

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`TERMINATED_WITH_ERRORS`	`CRITICAL`	Cambio de estado del clúster de EMR	ninguno	El clúster de Amazon EMR `ClusterId (ClusterName)` terminó con errores a las `Time` con el siguiente motivo: `StateChangeReason: Code` .

Eventos de cambio de estado de la flota de instancias

> **Note**
> La configuración de las flotas de instancias está disponible solo en las versiones 4.8.0 y posteriores de Amazon EMR, excluidas las versiones 5.0.0 y 5.0.3.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
Desde `PROVISIONING` hasta `WAITING`	`INFO`		ninguno	El aprovisionamiento para la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId (ClusterName)` está

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				completo. El aprovisio namiento se inició a las `Time` y tardó `Num` minutos. La flota de instancia s ahora tiene una capacidad bajo demanda de `Num` y una capacidad de spot de `Num`. La capacidad bajo demanda de destino era `Num`, y la capacidad de spot de destino era `Num`.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
Desde `WAITING` hasta `RESIZING`	`INFO`		ninguno	Se inició un cambio de tamaño de la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` se inició a las `Time`. La flota de instancias está cambiando de tamaño desde una capacidad bajo demanda de `Num` a un destino de `Num`, y de la capacidad de spot de `Num` a un destino de `Num`.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
Desde `RESIZING` hasta `WAITING`	`INFO`		ninguno	La operación de cambio de tamaño de la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` se ha completado. El cambio de tamaño se inició a las `Time` y tardó `Num` minutos. La flota de instancias ahora tiene una capacidad bajo demanda de `Num` y una capacidad de spot de `Num`. La capacidad bajo demanda de destino era `Num` y la capacidad de spot de destino era `Num`.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
Desde `RESIZING` hasta `WAITING`	`INFO`		ninguno	La operación de cambio de tamaño de la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` ha alcanzado el tiempo de espera y se ha parado. El cambio de tamaño se inició a las `Time` y se paró después de `Num` minutos. La flota de instancias ahora tiene una capacidad bajo demanda de `Num` y una capacidad de spot de `Num`. La capacidad bajo demanda de destino era `Num` y la capacidad de spot de destino era `Num`.

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`SUSPENDED`	`ERROR`		ninguno	La flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId (ClusterName)` se detuvo a las `Time` por el siguiente motivo: `ReasonDesc` .
`RESIZING`	`WARNING`		ninguno	La operación de cambio de tamaño de la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId (ClusterName)` está bloqueada por el siguiente motivo: `ReasonDesc` .

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
`WAITING` o `Running`	`INFO`		ninguno	La operación de redimensionamiento de la flota de instancias `InstanceFleetID` en el clúster de Amazon EMR `ClusterId` `(ClusterName)` no se pudo completar mientras Amazon EMR agregaba capacidad de spot en la zona de disponibilidad `AvailabilityZone`. Hemos cancelado su solicitud de aprovisionamiento de capacidad de spot adicional. Para ver las acciones recomendadas, consulte Flexibilidad de zona de disponibilidad

Estado o cambio de estado	Gravedad	Tipo de evento	Código de evento	Mensaje
				para un clúster de Amazon EMR e inténtelo de nuevo.
`WAITING` o `Running`	`INFO`		ninguno	`Entity` inició una operación de cambio de tamaño de la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` a las `Time`.

Eventos de reconfiguración de la flota de instancias

Estado o cambio de estado	Gravedad	Mensaje
Se ha solicitado la reconfiguración de la flota de instancias	`INFO`	Un usuario ha solicitado volver a configurar la flota de instancias `InstanceFleetID` en el `ClusterId` clúster `ClusterName` de Amazon EMR ().
Inicio de la reconfiguración de la flota de instancias	`INFO`	Amazon EMR ha iniciado una reconfiguración de la flota `InstanceFleetID` de instancias en el clúster

Estado o cambio de estado	Gravedad	Mensaje
		de Amazon EMR () en. `ClusterId` `ClusterName` `Time`
Se completó la reconfiguración de la flota de instancias	`INFO`	Amazon EMR ha terminado de reconfigurar la flota de instancias en `InstanceFleetID` el clúster de Amazon EMR (). `ClusterId` `ClusterName`
Falló la reconfiguración de la flota de instancias	`WARNING`	Amazon EMR no pudo reconfigurar la flota `InstanceFleetID` de instancias en el clúster de Amazon EMR () en. `ClusterId` `ClusterName` `Time` Se produjo un error en la reconfiguración debido a que. `Reason`
Inicio de la reversión de la reconfiguración de la flota de instancias	`INFO`	Amazon EMR está revirtiendo la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId` (`ClusterName`) a la configuración anterior correcta.

Estado o cambio de estado	Gravedad	Mensaje
Se completó la reconfiguración de la flota de instancias	`INFO`	Amazon EMR terminó de revertir la flota `InstanceFleetID` de instancias del clúster de Amazon EMR `ClusterId` (`ClusterName`) a la configuración anterior correcta.
Falló la reconfiguración de la flota de instancias	`CRITICAL`	Amazon EMR no pudo revertir la flota `InstanceFleetID` de instancias del clúster de Amazon EMR `ClusterId` (`ClusterName`) a la configuración anterior correcta en. `Time` La reversión de la reconfiguración falló debido a. `Reason`
Se bloqueó la reversión de la reconfiguración de la flota de instancias	`INFO`	Amazon EMR bloqueó temporalmente la flota de instancias en el clúster de `InstanceFleetID` Amazon EMR en `Time` porque la flota de `ClusterId` instancias `ClusterName` está en ese estado. `State`

Eventos de cambio de tamaño de la flota de instancias

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`ERROR`	Tiempo de espera de aprovisionamiento de spot	La operación de cambio de tamaño de la flota de instancia

Tipo de evento	Gravedad	Código de evento	Mensaje
			s `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` no se pudo completar al adquirir capacidad de spot en la zona de disponibilidad `AvailabilityZone` . Hemos cancelado su solicitud y hemos dejado de intentar aprovisionar capacidad de spot adicional. La flota de instancias tiene una capacidad de spot aprovisionada de `num`. La capacidad de spot de destino era de `num`. Para obtener más información y ver las acciones recomendadas, consulte la página de documentación aquí e inténtelo de nuevo.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`ERROR`	Tiempo de espera del aprovisionamiento bajo demanda	La operación de cambio de tamaño de la flota de instancias `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` no se pudo completar al adquirir capacidad bajo demanda en la zona de disponibilidad `AvailabilityZone`. Hemos cancelado su solicitud y hemos dejado de intentar aprovisionar capacidad bajo demanda adicional. La flota de instancias tiene una capacidad bajo demanda aprovisionada de `num`. La capacidad bajo demanda de destino era de `num`. Para obtener más información y ver las acciones recomendadas, consulte la página de documentación aquí e inténtelo de nuevo.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`WARNING`	EC2 aprovisio namiento: capacidad de instancia insuficie nte	No podemos completar la operación de cambio de tamaño de la flota de instancias `InstanceFleetID` en el clúster de EMR, `ClusterId (Cluster Name)` ya que EC2 Amazon no tiene suficiente capacidad puntual para los tipos de instancia s ni capacidad bajo demanda suficiente para los `[Instancetype1, Instancetype2]` tipos de instancia s en la zona de `[Instancetype3, Instancet ype4]` disponibi lidad. `[Availabi lityZone1]` Hasta ahora, la flota de instancias tiene una capacidad bajo demanda aprovisio nada de `num` y la capacidad bajo demanda de destino era de `num`. La

Tipo de evento	Gravedad	Código de evento	Mensaje
			capacidad de spot aprovisionada es de `num` y la capacidad de spot de destino es de `num`. Consulte aquí la documentación para obtener más información sobre cómo responder a este evento.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`WARNING`	Tiempo de espera de aprovisionamiento de spot: continuación del cambio de tamaño	Seguimos aprovisionando capacidad de spot para la operación de cambio de tamaño de la flota de instancias que se inició a las `time` en la flota de instancias con el ID `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` para `[Instancetype1, Instancetype2]` en la zona de disponibilidad `AvailabilityZone`. En la operación de redimensionamiento anterior que se inició a las `time`, se agotó el tiempo de espera, por lo que Amazon EMR dejó de aprovisionar la capacidad de spot tras agregar `num` de las `num` instancias solicitadas a su flota de instancias. Para obtener más información, consulte la página

Tipo de evento	Gravedad	Código de evento	Mensaje
			de documentación aquí.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`WARNING`	Tiempo de espera de aprovisionamiento bajo demanda: continuación del cambio de tamaño	Seguimos aprovisionando capacidad bajo demanda para la operación de cambio de tamaño de la flota de instancias que se inició a las `time` en la flota de instancias con el ID `InstanceFleetID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` para `[Instancetype1, Instancetype2]` en la zona de disponibilidad `AvailabilityZone`. En la operación de redimensionamiento anterior que se inició a las `time`, se agotó el tiempo de espera, por lo que Amazon EMR dejó de aprovisionar la capacidad bajo demanda tras agregar `num` de las `num` instancias solicitadas a su flota de instancias. Para obtener más informaci

Tipo de evento	Gravedad	Código de evento	Mensaje
			ón, consulte la página de documentación aquí.
Cambio de tamaño de la flota de instancias de EMR	`WARNING`	EC2 Aprovisio namiento: dirección libre insuficiente en la subred	No podemos completar la operación de cambio de tamaño para la flota de instancia s `InstanceF leetID` en el clúster de Amazon EMR `ClusterId (ClusterName)` porque la subred especificada [Subnet1, Subnet2] no contiene suficient es direcciones IP privadas libres para satisfacer su solicitud . Utilice la operación `DescribeSubnets` para ver cuántas direcciones IP están disponibles (no se utilizan) en su subred. Para obtener información sobre cómo responder a este evento, consulta Códigos de error de la EC2 API de Amazon.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`WARNING`	EC2 Aprovisio namiento: se ha superado el límite de vCPU	El cambio de tamaño de la flota de instancia s `InstanceF leetID` en el `ClusterName` clúster de Amazon EMR se retrasa porque ha alcanzado el límite de CPUs v (unidades de procesamiento virtuales) asignadas a las instancia s en ejecución en su. `account (accountId)` Para obtener más informaci ón, consulta Códigos de error de la EC2 API de Amazon.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`WARNING`	EC2 Aprovisio namiento: se ha superado el límite de número de instancias puntuales	El aprovisionamiento de la flota de instancia s `InstanceF leetID` en el clúster de Amazon EMR `ClusterID (ClusterName)` se retrasa porque ha alcanzado el límite del número de instancias de Spot que usted puede lanzar en su `account (accountId)` . Para obtener más informaci ón, consulta Códigos de error de la EC2 API de Amazon.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño de la flota de instancias de EMR	`WARNING`	EC2 Aprovisio namiento: se ha superado el límite de instancias	El aprovisionamiento de la flota de instancia s `InstanceF leetID` en el clúster de Amazon EMR `ClusterID (ClusterName)` se retrasa porque ha alcanzado el límite del número de instancia s bajo demanda que usted puede ejecutar en su `account (accountId)` . Para obtener más informaci ón sobre los códigos de error de la EC2 API de Amazon.

> **Note**
>
> Los eventos de tiempo de espera de aprovisionamiento se emiten cuando Amazon EMR deja de aprovisionar capacidad de spot o bajo demanda para la flota una vez transcurrido el tiempo de espera. Para obtener información sobre cómo responder a estos eventos, consulte Respuesta a eventos de tiempo de espera agotado en el cambio de tamaño de la flota de instancias de clústeres de Amazon EMR.

Eventos de grupo de instancias

Tipo de evento	Gravedad	Código de evento	Mensaje
Desde `RESIZING` hasta `Running`	`INFO`	ninguno	La operación de cambio de tamaño

Tipo de evento	Gravedad	Código de evento	Mensaje
			del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` se ha completado. Ahora tiene un recuento de instancias de `Num`. El cambio de tamaño se inició a las `Time` y tardó `Num` minutos en completarse.
Desde `RUNNING` hasta `RESIZING`	`INFO`	ninguno	Se inició un cambio de tamaño del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` a las `Time`. Está cambiando de tamaño de un recuento de instancias de `Num` a `Num`.

Tipo de evento	Gravedad	Código de evento	Mensaje
`SUSPENDED`	`ERROR`	ninguno	El grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` se detuvo a las `Time` por el siguiente motivo: `ReasonDesc` .
`RESIZING`	`WARNING`	ninguno	La operación de cambio de tamaño del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` está bloqueada por el siguiente motivo: `ReasonDesc` .

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño del grupo de instancias de EMR	`WARNING`	EC2 aprovisionamiento: capacidad de instancia insuficiente	No podemos completar la operación de cambio de tamaño que se inició en `time` el grupo `InstanceGroupID` de instancias del clúster de EMR `ClusterId (ClusterName)` porque EC2 Amazon `Spot/On Demand` no tiene suficiente capacidad para el `[Instancetype]` tipo de instancia en la zona de disponibilidad. `[AvailabilityZone1]` Hasta ahora, el grupo de instancias tiene un recuento de instancias de `num` y el número de instancias solicitadas era de `num`. Consulte aquí la documentación para obtener más información sobre cómo responder a este evento.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño del grupo de instancias de EMR	`WARNING`	EC2 Aprovisionamiento: dirección libre insuficiente en la subred	No podemos completar la operación de cambio de tamaño del grupo de instancias `InstanceGroupID` en el clúster de Amazon EMR `ClusterId (ClusterName)` porque la subred especificada [Subnet1, Subnet2] no contiene suficientes direcciones IP privadas libres para atender tu solicitud. Utilice la operación `DescribeSubnets` para ver cuántas direcciones IP están disponibles (no se utilizan) en su subred. Para obtener información sobre cómo responder a este evento, consulta Códigos de error de la EC2 API de Amazon.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño del grupo de instancias de EMR	`WARNING`	EC2 Aprovisionamiento: se ha superado el límite de vCPU	El cambio de tamaño del grupo de instancias `InstanceGroupID` en el `ClusterName` clúster de Amazon EMR se retrasa porque has alcanzado el límite de CPUs v (unidades de procesamiento virtuales) asignadas a las instancias en ejecución de tu. `account (accountId)` Para obtener más información, consulta Códigos de error de la EC2 API de Amazon.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño del grupo de instancias de EMR	`WARNING`	EC2 Aprovisionamiento: se ha superado el límite de número de instancias puntuales	El aprovisionamiento del grupo de instancias `InstanceGroupID` en el clúster de Amazon EMR `ClusterID (ClusterName)` se retrasa porque ha alcanzado el límite del número de instancias de Spot que puede lanzar en su `account (accountId)` . Para obtener más información, consulta Códigos de error de la EC2 API de Amazon.

Tipo de evento	Gravedad	Código de evento	Mensaje
Cambio de tamaño del grupo de instancias de EMR	`WARNING`	EC2 Aprovisionamiento: se ha superado el límite de instancias	El aprovisionamiento del grupo de instancias `InstanceGroupID` en el clúster de Amazon EMR `ClusterID (ClusterName)` se retrasa porque ha alcanzado el límite de instancias bajo demanda que puede ejecutar en su `account (accountId)`. Para obtener más información sobre los códigos de error de la EC2 API de Amazon.
Desde `RUNNING` hasta `RESIZING`	`INFO`	ninguno	`Entity` inició un cambio de tamaño del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId (ClusterName)` a las `Time`.

> **Note**
>
> Con la versión 5.21.0 y posteriores de Amazon EMR, puede anular las configuraciones de clúster y especificar las clasificaciones de configuración adicionales para cada grupo de instancias en un clúster en ejecución. Para ello, utilice la consola Amazon EMR, el AWS

Command Line Interface (AWS CLI) o el AWS SDK. Para obtener más información, consulte Suministrar una configuración para un grupo de instancias en un clúster en ejecución.

En la siguiente tabla se muestran eventos de Amazon EMR, junto con el estado o cambio de estado que indica el evento, la gravedad del evento y los mensajes de eventos.

Estado o cambio de estado	Gravedad	Mensaje
`RUNNING`	`INFO`	Un usuario inició un cambio de configuración del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` a las `Time`. La versión de configuración solicitada es `Num`.
Desde `RECONFIGURING` hasta `Running`	`INFO`	La operación de cambio de configuración del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` se ha completado. El cambio de configuración se inició a las `Time` y tardó `Num` minutos en completarse. La versión de la configuración actual es `Num`.
Desde `RUNNING` hasta `RECONFIGURING` in	`INFO`	Se inició un cambio de configuración del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` a las `Time`. Se configura de la

Estado o cambio de estado	Gravedad	Mensaje
		versión número Num a la versión número Num.
RESIZING	INFO	La operación de cambio de configuración hacia la versión de configuración Num para el grupo de instancias InstanceGroupID del clúster de Amazon EMR ClusterId (ClusterName) se ha bloqueado temporalmente a las Time porque el grupo de instancias se encuentra en State.
RECONFIGURING	INFO	La operación de cambio de tamaño hacia el recuento de instancias Num para el grupo de instancias InstanceGroupID del clúster de Amazon EMR ClusterId (ClusterName) se ha bloqueado temporalmente a las Time porque el grupo de instancias se encuentra en State.

Estado o cambio de estado	Gravedad	Mensaje
RECONFIGURING	WARNING	Se produjo un error en la operación de cambio de configuración del grupo de instancias InstanceGroupID del clúster de Amazon EMR ClusterId (ClusterName) a las Time y tardó Num minutos en producirse el error. La versión de configuración de error es Num.
RECONFIGURING	INFO	Las configuraciones se revierten al número de versión correcta anterior Num para el grupo de instancias InstanceGroupID del clúster de Amazon EMR ClusterId (ClusterName) a las Time. La nueva versión de configuración es Num.
Desde RECONFIGURING hasta Running	INFO	Las configuraciones se han revertido correctamente a la versión correcta anterior Num para el grupo de instancias InstanceGroupID del clúster de Amazon EMR ClusterId (ClusterName) a las Time. La nueva versión de configuración es Num.

Estado o cambio de estado	Gravedad	Mensaje
Desde `RECONFIGURING` hasta `SUSPENDED`	`CRITICAL`	No se pudo revertir a la versión correcta anterior `Num` para el grupo de instancia s `InstanceGroupID` del clúster de Amazon EMR `ClusterId (ClusterN ame)` a las `Time`.

Eventos de la política de escalado automático

Estado o cambio de estado	Gravedad	Mensaje
`PENDING`	`INFO`	Se ha agregado una política de escalado automátic o al grupo de instancias `InstanceGroupID` en el clúster de Amazon EMR `ClusterId (Cluster Name)` a las `Time`. La política todavía no se ha asociado. -o bien- Se ha actualizado la política de escalado automático del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId (Cluster Name)` a las `Time`. La política todavía no se ha asociado.

Estado o cambio de estado	Gravedad	Mensaje
`ATTACHED`	`INFO`	Se ha adjuntado la política de escalado automático del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` a las `Time`.
`DETACHED`	`INFO`	Se ha separado la política de escalado automático del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` a las `Time`.
`FAILED`	`ERROR`	No se pudo adjuntar la política de escalado automático del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` y se produjo un error a las `Time`. -o bien- No se pudo separar la política de escalado automático del grupo de instancias `InstanceGroupID` del clúster de Amazon EMR `ClusterId` `(ClusterName)` y se produjo un error a las `Time`.

Eventos de paso

Estado o cambio de estado	Gravedad	Mensaje
`PENDING`	`INFO`	El paso `StepID (StepName)` se agregó al clúster de Amazon EMR `ClusterId (ClusterName)` a las `Time` y está pendiente de ejecución.
`CANCEL_PENDING`	`WARN`	El paso `StepID (StepName)` del clúster de Amazon EMR `ClusterId (ClusterName)` se canceló a las `Time` y está pendiente de cancelación.
`RUNNING`	`INFO`	El paso `StepID (StepName)` del clúster de Amazon EMR `ClusterId (ClusterName)` comenzó a ejecutarse a las `Time`.
`COMPLETED`	`INFO`	El paso `StepID (StepName)` del clúster de Amazon EMR `ClusterId (ClusterName)` completó la ejecución a las `Time`. El paso empezó a ejecutarse a las `Time` y tardó `Num` minutos en completarse.
`CANCELLED`	`WARN`	La solicitud de cancelación se ha realizado correctam ente para el paso de clúster `StepID (StepName)` del clúster de Amazon EMR `ClusterId (Cluster`

Estado o cambio de estado	Gravedad	Mensaje
		`Name)` a las `Time`, y el paso ya está cancelado.
`FAILED`	`ERROR`	Se produjo un error en el paso `StepID (StepName)` del clúster de Amazon EMR `ClusterId (ClusterName)` a las `Time`.

Eventos de sustitución de nodos en mal estado

Tipo de evento	Gravedad	Código de evento	Mensaje	
Sustitución de nodo en mal estado en Amazon EMR	`INFO`	Se detectó un nodo básico en mal estado	Amazon EMR ha identificado la instancia básica `[instanceID (InstanceName)]` en `InstanceGroup/Fleet` en el clúster de Amazon EMR. `clusterID (ClusterName)` está en `UNHEALTHY`. Amazon EMR intentará recuperar o sustituir correctamente	

Tipo de evento	Gravedad	Código de evento	Mensaje	
			la instancia `UNHEALTHY` .	
Sustitución de nodo en mal estado en Amazon EMR	`INFO`	El nodo básico está en mal estado: el reemplazo se ha deshabilitado	Amazon EMR ha identificado la instancia básica `[instanceID (Instance Name)]` en `InstanceGroup/Fleet` en el clúster de Amazon EMR. `(clusterID) (ClusterName)` está en `UNHEALTHY` . Active la sustitución correcta del nodo básico en mal estado en su clúster para permitir que Amazon EMR sustituya correctamente las instancias `UNHEALTHY` en caso de que no se puedan recuperar.	

Tipo de evento	Gravedad	Código de evento	Mensaje	
Sustitución de nodo en mal estado en Amazon EMR	`WARN`	No se reemplazó el nodo básico en mal estado	Amazon EMR no puede reemplazar su instancia básica `[instanceID (Instance Name)]` `UNHEALTHY` en `InstanceGroup/Fleet` en el clúster de Amazon EMR `clusterID (ClusterName)` por algún motivo. **Note** El motivo por el que Amazon EMR no puede reemplazar su nodo básico es diferente según su escenario. Por ejemplo,	

Tipo de evento	Gravedad	Código de evento	Mensaje	
			uno de los motivos por los que Amazon EMR no puede eliminar un nodo es porque un clúster no tendría ningún nodo básico restante.	

Tipo de evento	Gravedad	Código de evento	Mensaje	
Sustitución de nodo en mal estado en Amazon EMR	`INFO`	Se ha recuperado un nodo básico en mal estado	Amazon EMR ha recuperado sus instancias básicas `UNHEALTHY` `[instanceID (Instance Name)]` en `InstanceGroup/Fleet` en el clúster de Amazon EMR `clusterID (ClusterName)`	

Para obtener más información sobre la sustitución de nodos en mal estado, consulte Sustitución de nodos en mal estado.

Visualización de eventos con la consola de Amazon EMR

Para cada clúster, puede ver una lista sencilla de eventos en el panel de detalles, que enumera los eventos en orden de aparición descendente. También puede ver todos los eventos de todos los clústeres en una región en orden de aparición descendente.

Si no desea que un usuario vea todos los eventos de clústeres de una región, añada una declaración que deniegue permiso (`"Effect": "Deny"`) para la acción `elasticmapreduce:ViewEventsFromAllClustersInConsole` a una política asociada al usuario.

Para ver los eventos de todos los clústeres de una región con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Eventos.

Para ver los eventos de un clúster determinado con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija un clúster.
3. Para ver todos los eventos, seleccione la pestaña Eventos en la página de detalles del clúster.

Respuesta a CloudWatch eventos de Amazon EMR

En esta sección, se describen varias formas en las que puede responder a eventos procesables que Amazon CloudWatch EMR emite como mensajes de eventos. Las formas de responder a los eventos incluyen la creación de reglas, la configuración de alarmas y otras respuestas. Las secciones siguientes incluyen enlaces a procedimientos y respuestas recomendadas a eventos comunes.

Temas

- Creación de reglas para eventos de Amazon EMR con CloudWatch
- Configuración de alarmas en CloudWatch las métricas de Amazon EMR
- Responder a eventos de capacidad de instancias insuficiente en el clúster de Amazon EMR
- Respuesta a eventos de tiempo de espera agotado en el cambio de tamaño de la flota de instancias de clústeres de Amazon EMR

Creación de reglas para eventos de Amazon EMR con CloudWatch

Amazon EMR envía automáticamente los eventos a una CloudWatch transmisión de eventos. Puede crear reglas que asignen eventos de acuerdo con un patrón especificado y dirijan los eventos a destinos para tomar medidas, por ejemplo, enviar una notificación por correo electrónico. Los patrones se asocian al objeto JSON del evento. Para obtener más información sobre los detalles de los eventos de Amazon EMR, consulte los eventos de Amazon EMR en la Guía del usuario de Amazon CloudWatch Events.

Para obtener información sobre cómo configurar las reglas de los CloudWatch eventos, consulte Crear una CloudWatch regla que se active en un evento.

Configuración de alarmas en CloudWatch las métricas de Amazon EMR

Amazon EMR envía las métricas a Amazon. CloudWatch En respuesta, puede utilizar CloudWatch para configurar alarmas en sus métricas de Amazon EMR. Por ejemplo, puede configurar una alarma CloudWatch para que le envíe un correo electrónico cada vez que la utilización del HDFS supere el 80%. Para obtener instrucciones detalladas, consulta Crear o editar una CloudWatch alarma en la Guía del CloudWatch usuario de Amazon.

Responder a eventos de capacidad de instancias insuficiente en el clúster de Amazon EMR

Descripción general

Los clústeres de Amazon EMR devuelven el código del evento `EC2 provisioning - Insufficient Instance Capacity` cuando la zona de disponibilidad seleccionada no tiene suficiente capacidad para cumplir con la solicitud de inicio o cambio de tamaño del clúster. El evento se emite periódicamente tanto con los grupos de instancias como con las flotas de instancias si Amazon EMR encuentra repetidamente excepciones de capacidad insuficiente y no puede cumplir con su solicitud de aprovisionamiento para una operación de inicio o cambio de tamaño del clúster.

En esta página, se describe la mejor manera de responder a este tipo de evento cuando se produce en su clúster de EMR.

Respuesta recomendada a un evento de capacidad insuficiente

Le recomendamos que responda a un evento de capacidad insuficiente de una de las siguientes maneras:

- Espere a que se recupere la capacidad. La capacidad cambia con frecuencia, por lo que una excepción de capacidad insuficiente puede recuperarse por sí sola. Tus clústeres empezarán o terminarán de redimensionarse en cuanto Amazon disponga de EC2 capacidad.
- Como alternativa, puede terminar el clúster, modificar las configuraciones del tipo de instancia y crear un nuevo clúster con la solicitud de configuración del clúster actualizada. Para obtener más información, consulte Flexibilidad de zona de disponibilidad para un clúster de Amazon EMR.

También puede configurar reglas o respuestas automatizadas a un evento de capacidad insuficiente, como se describe en la siguiente sección.

Recuperación automática de un evento de capacidad insuficiente

Puede crear una automatización en respuesta a los eventos de Amazon EMR, como los que tienen el código de evento `EC2 provisioning - Insufficient Instance Capacity`. Por ejemplo, la siguiente AWS Lambda función termina un clúster de EMR con un grupo de instancias que usa instancias bajo demanda y, a continuación, crea un nuevo clúster de EMR con un grupo de instancias que contiene tipos de instancias diferentes a los de la solicitud original.

Las siguientes condiciones activan el proceso automatizado:

- El evento de capacidad insuficiente se ha estado emitiendo en los nodos principales o de núcleo durante más de 20 minutos.
- El clúster no está en estado LISTO o EN ESPERA. Para obtener más información acerca de los estados del clúster de EMR, consulte Descripción del ciclo de vida del clúster.

> **Note**
>
> Al crear un proceso automatizado para una excepción de capacidad insuficiente, debe tener en cuenta que el evento de capacidad insuficiente es recuperable. La capacidad suele cambiar y tus clústeres reanudarán el cambio de tamaño o comenzarán a funcionar tan pronto como la EC2 capacidad de Amazon esté disponible.

Example función para responder a un evento de capacidad insuficiente

```
// Lambda code with Python 3.10 and handler is lambda_function.lambda_handler
// Note: related IAM role requires permission to use Amazon EMR

import json
import boto3
import datetime
from datetime import timezone

INSUFFICIENT_CAPACITY_EXCEPTION_DETAIL_TYPE = "EMR Instance Group Provisioning"
INSUFFICIENT_CAPACITY_EXCEPTION_EVENT_CODE = (
    "EC2 provisioning - Insufficient Instance Capacity"
)
ALLOWED_INSTANCE_TYPES_TO_USE = [
    "m5.xlarge",
    "c5.xlarge",
```

```
    "m5.4xlarge",
    "m5.2xlarge",
    "t3.xlarge",
]
CLUSTER_START_ACCEPTABLE_STATES = ["WAITING", "RUNNING"]
CLUSTER_START_SLA = 20

CLIENT = boto3.client("emr", region_name="us-east-1")

# checks if the incoming event is 'EMR Instance Fleet Provisioning' with eventCode 'EC2
 provisioning - Insufficient Instance Capacity'
def is_insufficient_capacity_event(event):
    if not event["detail"]:
        return False
    else:
        return (
            event["detail-type"] == INSUFFICIENT_CAPACITY_EXCEPTION_DETAIL_TYPE
            and event["detail"]["eventCode"]
            == INSUFFICIENT_CAPACITY_EXCEPTION_EVENT_CODE
        )


# checks if the cluster is eligible for termination
def is_cluster_eligible_for_termination(event, describeClusterResponse):
    # instanceGroupType could be CORE, MASTER OR TASK
    instanceGroupType = event["detail"]["instanceGroupType"]
    clusterCreationTime = describeClusterResponse["Cluster"]["Status"]["Timeline"][
        "CreationDateTime"
    ]
    clusterState = describeClusterResponse["Cluster"]["Status"]["State"]

    now = datetime.datetime.now()
    now = now.replace(tzinfo=timezone.utc)
    isClusterStartSlaBreached = clusterCreationTime < now - datetime.timedelta(
        minutes=CLUSTER_START_SLA
    )

    # Check if instance group receiving Insufficient capacity exception is CORE or
 PRIMARY (MASTER),
    # and it's been more than 20 minutes since cluster was created but the cluster
 state and the cluster state is not updated to RUNNING or WAITING
    if (
        (instanceGroupType == "CORE" or instanceGroupType == "MASTER")
        and isClusterStartSlaBreached
```

```
        and clusterState not in CLUSTER_START_ACCEPTABLE_STATES
    ):
        return True
    else:
        return False


# Choose item from the list except the exempt value
def choice_excluding(exempt):
    for i in ALLOWED_INSTANCE_TYPES_TO_USE:
        if i != exempt:
            return i


# Create a new cluster by choosing different InstanceType.
def create_cluster(event):
    # instanceGroupType cloud be CORE, MASTER OR TASK
    instanceGroupType = event["detail"]["instanceGroupType"]

    # Following two lines assumes that the customer that created the cluster already
 knows which instance types they use in original request
    instanceTypesFromOriginalRequestMaster = "m5.xlarge"
    instanceTypesFromOriginalRequestCore = "m5.xlarge"

    # Select new instance types to include in the new createCluster request
    instanceTypeForMaster = (
        instanceTypesFromOriginalRequestMaster
        if instanceGroupType != "MASTER"
        else choice_excluding(instanceTypesFromOriginalRequestMaster)
    )
    instanceTypeForCore = (
        instanceTypesFromOriginalRequestCore
        if instanceGroupType != "CORE"
        else choice_excluding(instanceTypesFromOriginalRequestCore)
    )

    print("Starting to create cluster...")
    instances = {
        "InstanceGroups": [
            {
                "InstanceRole": "MASTER",
                "InstanceCount": 1,
                "InstanceType": instanceTypeForMaster,
                "Market": "ON_DEMAND",
```

```
                "Name": "Master",
            },
            {
                "InstanceRole": "CORE",
                "InstanceCount": 1,
                "InstanceType": instanceTypeForCore,
                "Market": "ON_DEMAND",
                "Name": "Core",
            },
        ]
    }
    response = CLIENT.run_job_flow(
        Name="Test Cluster",
        Instances=instances,
        VisibleToAllUsers=True,
        JobFlowRole="EMR_EC2_DefaultRole",
        ServiceRole="EMR_DefaultRole",
        ReleaseLabel="emr-6.10.0",
    )

    return response["JobFlowId"]


# Terminated the cluster using clusterId received in an event
def terminate_cluster(event):
    print("Trying to terminate cluster, clusterId: " + event["detail"]["clusterId"])
    response = CLIENT.terminate_job_flows(JobFlowIds=[event["detail"]["clusterId"]])
    print(f"Terminate cluster response: {response}")


def describe_cluster(event):
    response = CLIENT.describe_cluster(ClusterId=event["detail"]["clusterId"])
    return response


def lambda_handler(event, context):
    if is_insufficient_capacity_event(event):
        print(
            "Received insufficient capacity event for instanceGroup, clusterId: "
            + event["detail"]["clusterId"]
        )

        describeClusterResponse = describe_cluster(event)
```

```
        shouldTerminateCluster = is_cluster_eligible_for_termination(
            event, describeClusterResponse
        )
        if shouldTerminateCluster:
            terminate_cluster(event)

            clusterId = create_cluster(event)
            print("Created a new cluster, clusterId: " + clusterId)
        else:
            print(
                "Cluster is not eligible for termination, clusterId: "
                + event["detail"]["clusterId"]
            )

    else:
        print("Received event is not insufficient capacity event, skipping")
```

Respuesta a eventos de tiempo de espera agotado en el cambio de tamaño de la flota de instancias de clústeres de Amazon EMR

Descripción general

Los clústeres de Amazon EMR emiten eventos al ejecutar la operación de cambio de tamaño en los clústeres de flotas de instancias. Los eventos de tiempo de espera de aprovisionamiento se emiten cuando Amazon EMR deja de aprovisionar capacidad de spot o bajo demanda para la flota una vez transcurrido el tiempo de espera. El usuario puede configurar la duración del tiempo de espera como parte de las especificaciones de cambio de tamaño de las flotas de instancias. En escenarios de cambios de tamaño consecutivos para la misma flota de instancias, Amazon EMR emite los eventos `Spot provisioning timeout - continuing resize` o `On-Demand provisioning timeout - continuing resize` cuando vence el tiempo de espera de la operación de cambio de tamaño actual. A continuación, comienza a aprovisionar capacidad para la siguiente operación de cambio de tamaño de la flota.

Respuesta a eventos de tiempo de espera agotado en el cambio de tamaño de la flota de instancias

Le recomendamos que responda a un evento de tiempo de espera de aprovisionamiento de una de las siguientes maneras:

- Revisite las especificaciones de cambio de tamaño y vuelva a intentar la operación de cambio de tamaño. Como la capacidad cambia con frecuencia, tus clústeres cambiarán de tamaño correctamente en cuanto Amazon disponga EC2 de capacidad. Recomendamos a los clientes que

configuren valores más bajos para el tiempo de espera para los trabajos que requieren un nivel más estricto. SLAs

- Como alternativa, puede:
  - Lance un nuevo clúster con tipos de instancias diversificados según las prácticas recomendadas para la flexibilidad de las instancias y las zonas de disponibilidad o
  - Lanzar un clúster con capacidad bajo demanda
- En el caso de un evento “Tiempo de espera de aprovisionamiento: continuación del cambio de tamaño”, también puede esperar a que se procesen las operaciones de cambio de tamaño. Amazon EMR seguirá procesando secuencialmente las operaciones de cambio de tamaño activadas para la flota, respetando las especificaciones de cambio de tamaño configuradas.

También puede configurar reglas o respuestas automatizadas a este evento, como se describe en la siguiente sección.

Recuperación automática de un evento de tiempo de espera de aprovisionamiento

Puede crear una automatización en respuesta a los eventos de Amazon EMR con el código de evento `Spot Provisioning timeout`. Por ejemplo, la siguiente función de AWS Lambda apaga un clúster de EMR con una flota de instancias que usa instancias de spot para los nodos de tarea y, a continuación, crea un nuevo clúster de EMR con una flota de instancias que contiene tipos de instancias más diversificados que la solicitud original. En este ejemplo, el evento `Spot Provisioning timeout` emitido para los nodos de tarea activará la ejecución de la función de Lambda.

Example Ejemplo de función para responder a un evento **`Spot Provisioning timeout`**

```
// Lambda code with Python 3.10 and handler is lambda_function.lambda_handler
// Note: related IAM role requires permission to use Amazon EMR

import json
import boto3
import datetime
from datetime import timezone

SPOT_PROVISIONING_TIMEOUT_EXCEPTION_DETAIL_TYPE = "EMR Instance Fleet Resize"
SPOT_PROVISIONING_TIMEOUT_EXCEPTION_EVENT_CODE = (
    "Spot Provisioning timeout"
)
```

```
CLIENT = boto3.client("emr", region_name="us-east-1")

# checks if the incoming event is 'EMR Instance Fleet Resize' with eventCode 'Spot
 provisioning timeout'
def is_spot_provisioning_timeout_event(event):
    if not event["detail"]:
        return False
    else:
        return (
            event["detail-type"] == SPOT_PROVISIONING_TIMEOUT_EXCEPTION_DETAIL_TYPE
            and event["detail"]["eventCode"]
            == SPOT_PROVISIONING_TIMEOUT_EXCEPTION_EVENT_CODE
        )


# checks if the cluster is eligible for termination
def is_cluster_eligible_for_termination(event, describeClusterResponse):
    # instanceFleetType could be CORE, MASTER OR TASK
    instanceFleetType = event["detail"]["instanceFleetType"]

    # Check if instance fleet receiving Spot provisioning timeout event is TASK
    if (instanceFleetType == "TASK"):
        return True
    else:
        return False


# create a new cluster by choosing different InstanceType.
def create_cluster(event):
    # instanceFleetType cloud be CORE, MASTER OR TASK
    instanceFleetType = event["detail"]["instanceFleetType"]

    # the following two lines assumes that the customer that created the cluster
 already knows which instance types they use in original request
    instanceTypesFromOriginalRequestMaster = "m5.xlarge"
    instanceTypesFromOriginalRequestCore = "m5.xlarge"

    # select new instance types to include in the new createCluster request
    instanceTypesForTask = [
        "m5.xlarge",
        "m5.2xlarge",
        "m5.4xlarge",
        "m5.8xlarge",
        "m5.12xlarge"
```

```
    ]

    print("Starting to create cluster...")
    instances = {
        "InstanceFleets": [
            {
                "InstanceFleetType":"MASTER",
                "TargetOnDemandCapacity":1,
                "TargetSpotCapacity":0,
                "InstanceTypeConfigs":[
                    {
                        'InstanceType': instanceTypesFromOriginalRequestMaster,
                        "WeightedCapacity":1,
                    }
                ]
            },
            {
                "InstanceFleetType":"CORE",
                "TargetOnDemandCapacity":1,
                "TargetSpotCapacity":0,
                "InstanceTypeConfigs":[
                    {
                        'InstanceType': instanceTypesFromOriginalRequestCore,
                        "WeightedCapacity":1,
                    }
                ]
            },
            {
                "InstanceFleetType":"TASK",
                "TargetOnDemandCapacity":0,
                "TargetSpotCapacity":100,
                "LaunchSpecifications":{},
                "InstanceTypeConfigs":[
                    {
                        'InstanceType': instanceTypesForTask[0],
                        "WeightedCapacity":1,
                    },
                    {
                        'InstanceType': instanceTypesForTask[1],
                        "WeightedCapacity":2,
                    },
                    {
                        'InstanceType': instanceTypesForTask[2],
                        "WeightedCapacity":4,
```

```
                    },
                    {
                        'InstanceType': instanceTypesForTask[3],
                        "WeightedCapacity":8,
                    },
                    {
                        'InstanceType': instanceTypesForTask[4],
                        "WeightedCapacity":12,
                    }
                ],
                "ResizeSpecifications": {
                    "SpotResizeSpecification": {
                        "TimeoutDurationMinutes": 30
                    }
                }
            }
        ]
    }
    response = CLIENT.run_job_flow(
        Name="Test Cluster",
        Instances=instances,
        VisibleToAllUsers=True,
        JobFlowRole="EMR_EC2_DefaultRole",
        ServiceRole="EMR_DefaultRole",
        ReleaseLabel="emr-6.10.0",
    )

    return response["JobFlowId"]


# terminated the cluster using clusterId received in an event
def terminate_cluster(event):
    print("Trying to terminate cluster, clusterId: " + event["detail"]["clusterId"])
    response = CLIENT.terminate_job_flows(JobFlowIds=[event["detail"]["clusterId"]])
    print(f"Terminate cluster response: {response}")


def describe_cluster(event):
    response = CLIENT.describe_cluster(ClusterId=event["detail"]["clusterId"])
    return response


def lambda_handler(event, context):
    if is_spot_provisioning_timeout_event(event):
```

```
        print(
            "Received spot provisioning timeout event for instanceFleet, clusterId: "
            + event["detail"]["clusterId"]
        )

        describeClusterResponse = describe_cluster(event)

        shouldTerminateCluster = is_cluster_eligible_for_termination(
            event, describeClusterResponse
        )
        if shouldTerminateCluster:
            terminate_cluster(event)

            clusterId = create_cluster(event)
            print("Created a new cluster, clusterId: " + clusterId)
        else:
            print(
                "Cluster is not eligible for termination, clusterId: "
                + event["detail"]["clusterId"]
            )

    else:
        print("Received event is not spot provisioning timeout event, skipping")
```

Visualización de métricas de aplicaciones de clúster mediante Ganglia con Amazon EMR

Ganglia está disponible con las versiones 4.2 a 6.15 de Amazon EMR. Ganglia es un proyecto de código que es un sistema distribuido y escalable diseñado para monitorizar clústeres y redes al mismo tiempo que minimiza el impacto en su rendimiento. Al habilitar Ganglia en su clúster, puede generar informes y ver el rendimiento del clúster en su conjunto, además de inspeccionar el rendimiento de cada una de las instancias de nodo individuales. Ganglia también está configurado para adquirir y visualizar las métricas de Hadoop y Spark. Para obtener más información, consulte Ganglia en la Guía de publicación de Amazon EMR.

Registro de llamadas a la API AWS EMR mediante AWS CloudTrail

AWS EMR está integrado con AWS CloudTrailun servicio que proporciona un registro de las acciones realizadas por un usuario, rol o un. Servicio de AWS CloudTrail captura todas las llamadas a la API para AWS EMR como eventos. Las llamadas capturadas incluyen llamadas desde la consola de AWS EMR y llamadas de código a las operaciones de la API de AWS EMR. Con la información

recopilada por CloudTrail, puede determinar la solicitud que se realizó a AWS EMR, la dirección IP desde la que se realizó la solicitud, cuándo se realizó y detalles adicionales.

Cada entrada de registro o evento contiene información sobre quién generó la solicitud. La información de identidad del usuario le ayuda a determinar lo siguiente:

- Si la solicitud se realizó con las credenciales del usuario raíz o del usuario.
- Si la solicitud se realizó en nombre de un usuario de IAM Identity Center.
- Si la solicitud se realizó con credenciales de seguridad temporales de un rol o fue un usuario federado.
- Si la solicitud la realizó otro Servicio de AWS.

CloudTrail está activa en tu cuenta Cuenta de AWS cuando creas la cuenta y tienes acceso automáticamente al historial de CloudTrail eventos. El historial de CloudTrail eventos proporciona un registro visible, consultable, descargable e inmutable de los últimos 90 días de eventos de gestión registrados en un. Región de AWSPara obtener más información, consulte Uso del historial de CloudTrail eventos en la Guía del usuario.AWS CloudTrail La visualización del historial de eventos no conlleva ningún CloudTrail cargo.

Para tener un registro continuo de los eventos de Cuenta de AWS los últimos 90 días, crea un almacén de datos de eventos de senderos o CloudTraillagos.

CloudTrail senderos

Un rastro permite CloudTrail entregar archivos de registro a un bucket de Amazon S3. Todos los senderos creados con él AWS Management Console son multirregionales. Puede crear un registro de seguimiento de una sola región o multirregionales mediante la AWS CLI. Se recomienda crear un sendero multirregional, ya que puedes capturar toda la actividad de tu Regiones de AWS cuenta. Si crea un registro de seguimiento de una sola región, solo podrá ver los eventos registrados en la Región de AWS del registro de seguimiento. Para obtener más información acerca de los registros de seguimiento, consulte Creación de un registro de seguimiento para su Cuenta de AWS y Creación de un registro de seguimiento para una organización en la Guía del usuario de AWS CloudTrail .

Puede enviar una copia de sus eventos de administración en curso a su bucket de Amazon S3 sin coste alguno CloudTrail mediante la creación de una ruta; sin embargo, hay cargos por almacenamiento en Amazon S3. Para obtener más información sobre CloudTrail los precios,

consulte AWS CloudTrail Precios. Para obtener información acerca de los precios de Amazon S3, consulte Precios de Amazon S3.

CloudTrail Almacenes de datos de eventos en Lake

CloudTrail Lake le permite ejecutar consultas basadas en SQL en sus eventos. CloudTrail Lake convierte los eventos existentes en formato JSON basado en filas al formato Apache ORC. ORC es un formato de almacenamiento en columnas optimizado para una recuperación rápida de datos. Los eventos se agregan en almacenes de datos de eventos, que son recopilaciones inmutables de eventos en función de criterios que se seleccionan aplicando selectores de eventos avanzados. Los selectores que se aplican a un almacén de datos de eventos controlan los eventos que perduran y están disponibles para la consulta. Para obtener más información sobre CloudTrail Lake, consulte Cómo trabajar con AWS CloudTrail Lake en la Guía del AWS CloudTrail usuario.

CloudTrail Los almacenes de datos y las consultas sobre eventos de Lake conllevan costes. Cuando crea un almacén de datos de eventos, debe elegir la opción de precios que desee utilizar para él. La opción de precios determina el costo de la incorporación y el almacenamiento de los eventos, así como el período de retención predeterminado y máximo del almacén de datos de eventos. Para obtener más información sobre CloudTrail los precios, consulte AWS CloudTrail Precios.

AWS Eventos de datos EMR en CloudTrail

Los eventos de datos proporcionan información sobre las operaciones de recursos realizadas en o dentro de un recurso (por ejemplo, leer o escribir en un objeto de Amazon S3). Se denominan también operaciones del plano de datos. Los eventos de datos suelen ser actividades de gran volumen. De forma predeterminada, CloudTrail no registra los eventos de datos. El historial de CloudTrail eventos no registra los eventos de datos.

Se aplican cargos adicionales a los eventos de datos. Para obtener más información sobre CloudTrail los precios, consulta AWS CloudTrail Precios.

Puede registrar eventos de datos para los tipos de recursos de AWS EMR mediante la CloudTrail consola o las operaciones AWS CLI de la CloudTrail API. Para obtener más información sobre cómo registrar los eventos de datos, consulte Registro de eventos de datos con la AWS Management Console y Registro de eventos de datos con la AWS Command Line Interface en la Guía del usuario de AWS CloudTrail .

En la siguiente tabla se enumeran los tipos de recursos de AWS EMR para los que puede registrar eventos de datos. La columna Tipo de evento de datos (consola) muestra el valor que se puede elegir en la lista de tipos de eventos de datos de la CloudTrail consola. La columna de valores resources.type muestra el `resources.type` valor, que se especificaría al configurar los selectores de eventos avanzados mediante o. AWS CLI CloudTrail APIs La CloudTrail columna Datos APIs registrados muestra las llamadas a la API registradas CloudTrail para el tipo de recurso.

Para obtener más información sobre las operaciones de la API, consulte la Referencia de la CLI de Amazon EMR WAL (EMRWAL). Amazon EMR registra algunas operaciones de la API de datos CloudTrail que son operaciones HBase del sistema a las que nunca se llama directamente. Estas operaciones no están en la referencia de la CLI de EMRWAL.

Tipo de evento de datos (consola)	resources.type value	Datos APIs registrados en CloudTrail
Espacio de trabajo de registro de escritura anticipada de Amazon EMR	`AWS::EMRWAL::Workspace`	• GetCurrentWALTime • ListTagsForResource • Lista WALs • ListWorkspaces • TrimWAL • Completa WALFlush

Puede configurar selectores de eventos avanzados para filtrar según los campos `eventName`, `readOnly` y `resources.ARN` y así registrar solo los eventos que son importantes para usted. Para obtener más información acerca de estos campos, consulte AdvancedFieldSelector en la Referencia de la API de AWS CloudTrail .

AWS Eventos de gestión de EMR en CloudTrail

Los eventos de administración proporcionan información sobre las operaciones de administración que se llevan a cabo en los recursos de su Cuenta de AWS empresa. Se denominan también operaciones del plano de control. De forma predeterminada, CloudTrail registra los eventos de administración.

AWS EMR registra todas las operaciones del plano de control de AWS EMR como eventos de administración. Para obtener una lista de las operaciones del plano de control de AWS EMR en las que se registra AWS EMR, CloudTrail consulte la referencia de la API de AWS EMR.

AWS Ejemplos de eventos EMR

Un evento representa una solicitud única de cualquier fuente e incluye información sobre la operación de API solicitada, la fecha y la hora de la operación, los parámetros de la solicitud, etc. CloudTrail Los archivos de registro no son un registro ordenado de las llamadas a la API pública, por lo que los eventos no aparecen en ningún orden específico.

En el siguiente ejemplo, se muestra una entrada de CloudTrail registro que demuestra la RunJobFlowacción.

```
{
 "Records": [
 {
         "eventVersion":"1.01",
         "userIdentity":{
            "type":"IAMUser",
            "principalId":"EX_PRINCIPAL_ID",
            "arn":"arn:aws:iam::123456789012:user/temporary-user-xx-7M",
            "accountId":"123456789012",
            "userName":"temporary-user-xx-7M"
         },
         "eventTime":"2018-03-31T17:59:21Z",
         "eventSource":"elasticmapreduce.amazonaws.com",
         "eventName":"RunJobFlow",
         "awsRegion":"us-west-2",
         "sourceIPAddress":"192.0.2.1",
         "userAgent":"aws-sdk-java/unknown-version Linux/xx Java_HotSpot(TM)_64-
Bit_Server_VM/xx",
         "requestParameters":{
            "tags":[
               {
                  "value":"prod",
                  "key":"domain"
               },
               {
                  "value":"us-west-2",
                  "key":"realm"
               },
               {
                  "value":"VERIFICATION",
                  "key":"executionType"
               }
```

```
            ],
            "instances":{
               "slaveInstanceType":"m5.xlarge",
               "ec2KeyName":"emr-integtest",
               "instanceCount":1,
               "masterInstanceType":"m5.xlarge",
               "keepJobFlowAliveWhenNoSteps":true,
               "terminationProtected":false
            },
            "visibleToAllUsers":false,
            "name":"MyCluster",
            "ReleaseLabel":"emr-5.16.0"
         },
         "responseElements":{
            "jobFlowId":"j-2WDJCGEG4E6AJ"
         },
         "requestID":"2f482daf-b8fe-11e3-89e7-75a3d0e071c5",
         "eventID":"b348a38d-f744-4097-8b2a-e68c9b424698"
      },
 ...additional entries
  ]
}
```

Para obtener información sobre el contenido de los CloudTrail registros, consulte el contenido de los CloudTrail registros en la Guía del AWS CloudTrail usuario.

Mejores prácticas de observabilidad de EMR

La observabilidad de EMR abarca un enfoque integral de monitoreo y gestión para los clústeres de AWS EMR. La base se basa en Amazon CloudWatch como servicio de monitoreo principal, complementado por EMR Studio y herramientas de terceros como Prometheus y Grafana para mejorar la visibilidad. En este documento, exploramos aspectos específicos de la observabilidad de los clústeres:

1. Observabilidad de Spark (GitHub): con respecto a la interfaz de usuario de Spark, tiene tres opciones en Amazon EMR.
2. Solución de problemas con Spark (GitHub): soluciones para errores.
3. Monitoreo de clústeres EMR (GitHub): monitoreo del rendimiento del clúster.
4. Solución de problemas de EMR (GitHub): identifique, diagnostique y resuelva problemas comunes del clúster de EMR.

5. Optimización de costos (GitHub): en esta sección se describen las mejores prácticas para ejecutar cargas de trabajo rentables.

Herramienta de optimización del rendimiento para aplicaciones de Apache Spark

1. AWS La herramienta EMR Advisor analiza los registros de eventos de Spark para ofrecer recomendaciones personalizadas para optimizar las configuraciones de los clústeres de EMR, mejorar el rendimiento y reducir los costos. Al aprovechar los datos históricos, sugiere tamaños de ejecutores y configuraciones de infraestructura ideales, lo que permite una utilización más eficiente de los recursos y un mejor rendimiento general del clúster.
2. La herramienta Amazon CodeGuru Profiler ayuda a los desarrolladores a identificar los cuellos de botella y las ineficiencias en el rendimiento de sus aplicaciones Spark mediante la recopilación y el análisis de los datos de tiempo de ejecución. La herramienta se integra perfectamente con las aplicaciones de Spark existentes, requiere una configuración mínima y proporciona información detallada a través de la AWS consola sobre el uso de la CPU, los patrones de memoria y los puntos críticos de rendimiento.

Utilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes

Puede ajustar el número de EC2 instancias de Amazon disponibles en un clúster de Amazon EMR de forma automática o manual en respuesta a cargas de trabajo que tengan diferentes demandas. Para utilizar el escalado automático, tiene dos opciones. Puede habilitar Escalado administrado de Amazon EMR o crear una política de escalado automático personalizada. En la siguiente tabla se describen las diferencias entre las dos opciones.

	Escalado administrado de Amazon EMR	Escalado automático personalizado
Políticas y reglas de escalado	No se requiere ninguna política. Amazon EMR administra la actividad de escalado automático evaluando continuamente las métricas del clúster y tomando	Debe definir y administrar las políticas y reglas de escalado automático, como las condiciones específicas que desencadenan actividades de escalado, periodos

	Escalado administrado de Amazon EMR	Escalado automático personalizado
	decisiones de escalado optimizado.	de evaluación, periodos de recuperación, etc.
Versiones compatibles de Amazon EMR	Versiones 5.30.0 y posteriores de Amazon EMR (excepto la versión 6.0.0 de Amazon EMR)	Versiones 4.0.0 y posteriores de Amazon EMR
Composición del clúster admitida	Grupos de instancias o flotas de instancias	Solo grupos de instancias
Configuración de límites de escalado	Los límites de escalado se configuran para todo el clúster.	Los límites de escalado solo se pueden configurar para cada grupo de instancias.
Frecuencia de evaluación de las métricas	Cada 5-10 segundos Una evaluación más frecuente de las métricas permite a Amazon EMR tomar decisiones de escalado más precisas.	Puede definir los periodos de evaluación solo en incrementos de cinco minutos.
Aplicaciones compatibles	Solo se admiten aplicaciones YARN, como Spark, Hadoop, Hive y Flink. El escalado gestionado por Amazon EMR no admite aplicaciones que no estén basadas en YARN, como Presto o. HBase	Puede elegir qué aplicaciones se admiten cuando defina las reglas de escalado automático.

Consideraciones

- Un clúster de Amazon EMR siempre consta de uno o tres nodos principales. Una vez que haya configurado inicialmente el clúster, solo podrá escalar los nodos principales y de tareas. No puede escalar el número de nodos principales del clúster.

- En el caso de los grupos de instancias, las operaciones de reconfiguración y cambio de tamaño se realizan de forma consecutiva y no simultánea. Si inicia una reconfiguración mientras se está cambiando el tamaño de un grupo de instancias, la reconfiguración comienza una vez que el grupo de instancias complete el cambio de tamaño en curso. Por el contrario, si inicia una operación de cambio de tamaño mientras se está reconfigurando un grupo de instancias, se cambiará el tamaño en el momento.

Uso del escalado administrado en Amazon EMR

> **Important**
>
> Le recomendamos encarecidamente que utilice la versión más reciente de Amazon EMR (Amazon EMR 7.10.0) para gestionar el escalado. En las versiones anteriores, era posible que se produjeran errores intermitentes en las aplicaciones o retrasos en el escalado. Amazon EMR resolvió este problema en las versiones 5.x: (5.30.2, 5.31.1, 5.32.1, 5.33.1 y posteriores) y las versiones 6.x (6.1.1, 6.2.1, 6.3.1 y posteriores). Para obtener más información sobre la disponibilidad de las versiones y las regiones, consulte Disponibilidad de Escalado administrado.

Descripción general

Con las versiones 5.30.0 y posteriores de Amazon EMR (excepto la versión 6.0.0 de Amazon EMR), puede habilitar el Escalado administrado de Amazon EMR. El escalado administrado le permite aumentar o disminuir automáticamente el número de instancias o unidades del clúster en función de la carga de trabajo. Amazon EMR evalúa continuamente las métricas del clúster para tomar decisiones de escalado que optimicen los clústeres en cuanto al costo y la velocidad. El escalado administrado está disponible para clústeres compuestos por grupos de instancias o flotas de instancias.

Disponibilidad de Escalado administrado

- A continuación Regiones de AWS, el escalado gestionado de Amazon EMR está disponible con Amazon EMR 6.14.0 y versiones posteriores:
  - Europa (España) (eu-south-2)
- A continuación Regiones de AWS, el escalado gestionado de Amazon EMR está disponible con Amazon EMR 5.30.0 y 6.1.0 y versiones posteriores:

- Este de EE. UU. (Norte de Virginia) (us-east-1)
- Este de EE. UU. (Ohio) (us-east-2)
- Oeste de EE. UU. (Oregón) (us-west-2)
- EE. UU. Oeste (Norte de California) (us-west-1)
- África (Ciudad del Cabo) (af-south-1)
- Asia-Pacífico (Hong Kong) (ap-east-1)
- Asia Pacífico (Bombay) (ap-south-1)
- Asia Pacífico (Hyderabad) (ap-south-2)
- Asia-Pacífico (Seúl) (ap-northeast-2)
- Asia-Pacífico (Singapur) (ap-southeast-1)
- Asia-Pacífico (Sídney) (ap-southeast-2)
- Asia-Pacífico (Yakarta) (ap-southeast-3)
- Asia-Pacífico (Tokio) (ap-northeast-1)
- Asia Pacific (Osaka) (ap-northeast-3)
- Canadá (centro) (ca-central-1)
- América del Sur (São Paulo) (sa-east-1)
- Europa (Fráncfort) (eu-central-1)
- Europa (Zúrich) (eu-central-2)
- Europa (Irlanda) (eu-west-1)
- Europa (Londres) (eu-west-2)
- UE (Milán) (eu-south-1)
- UE (París) (eu-west-3)
- Europa (Estocolmo) (eu-north-1)
- Israel (Tel Aviv) (il-central-1)
- Medio Oriente (EAU) (me-central-1)
- China (Pekín) (cn-north-1)
- China (Ningxia) (cn-northwest-1)
- AWS GovCloud (EE. UU.-Este) (-1) us-gov-east
- AWS GovCloud (EEUU-Oeste) (us-gov-west-1)

- Escalado administrado de Amazon EMR solo funciona con aplicaciones YARN, como Spark, Hadoop, Hive y Flink. No admite aplicaciones que no estén basadas en YARN, como Presto y. HBase

Parámetros de escalado administrado

Debe configurar los siguientes parámetros para el escalado administrado. El límite solo se aplica a los nodos principales y de tareas. No puede escalar el nodo principal después de la configuración inicial.

- Mínimo (`MinimumCapacityUnits`): el límite inferior de la EC2 capacidad permitida en un clúster. Se mide mediante núcleos de unidades de procesamiento central virtual (vCPU) o instancias para grupos de instancias. Se mide mediante unidades para flotas de instancias.
- Máximo (`MaximumCapacityUnits`): el límite superior de la EC2 capacidad permitida en un clúster. Se mide mediante núcleos de unidades de procesamiento central virtual (vCPU) o instancias para grupos de instancias. Se mide mediante unidades para flotas de instancias.
- Límite bajo demanda (`MaximumOnDemandCapacityUnits`) (opcional): límite superior de la EC2 capacidad permitida para el tipo de mercado bajo demanda en un clúster. Si no se especifica este parámetro, se establece en el valor predeterminado de `MaximumCapacityUnits`.
  - Este parámetro se utiliza para dividir la asignación de capacidad entre instancias bajo demanda e instancias de spot. Por ejemplo, si establece el parámetro mínimo en 2 instancias, el parámetro máximo en 100 instancias y el límite bajo demanda en 10 instancias, Escalado administrado de Amazon EMR escala hasta 10 instancias bajo demanda y asigna la capacidad restante a instancias de spot. Para obtener más información, consulte Escenarios de asignación de nodos.
- Número máximo de nodos principales (`MaximumCoreCapacityUnits`) (opcional): límite superior de la EC2 capacidad permitida para el tipo de nodo principal de un clúster. Si no se especifica este parámetro, se establece en el valor predeterminado de `MaximumCapacityUnits`.
  - Este parámetro se utiliza para dividir la asignación de capacidad entre los nodos principales y de tarea. Por ejemplo, si establece el parámetro mínimo en 2 instancias, el máximo en 100 instancias y el máximo de nodos principales en 17 instancias, Escalado administrado de Amazon EMR escala hasta 17 nodos principales y asigna las 83 instancias restantes a los nodos de tarea. Para obtener más información, consulte Escenarios de asignación de nodos.

Para obtener más información sobre los parámetros de escalado administrado, consulte `ComputeLimits`.

Consideraciones sobre el Escalado administrado de Amazon EMR

- El escalado gestionado se admite en las versiones limitadas Regiones de AWS y de Amazon EMR. Para obtener más información, consulte Disponibilidad de Escalado administrado.
- Debe configurar los parámetros obligatorios para Escalado administrado de Amazon EMR. Para obtener más información, consulte Parámetros de escalado administrado.
- Para utilizar el escalado administrado, el proceso de recopilación de métricas debe poder conectarse al punto de conexión de la API pública para el escalado administrado en API Gateway. Si utiliza un nombre de DNS privado con Amazon Virtual Private Cloud, el escalado administrado no funcionará correctamente. Para garantizar que el escalado administrado funcione, se recomienda que realice una de las siguientes acciones:
  - Elimine el punto de conexión de VPC de la interfaz de API Gateway de su Amazon VPC.
  - Siga las instrucciones de ¿Por qué aparece el error HTTP 403 Forbidden al conectarme a mi API Gateway APIs desde una VPC? para deshabilitar la configuración del nombre DNS privado.
  - En su lugar, lance su clúster en una subred privada. Para más información, consulte el tema sobre Subredes privadas.
- Si sus trabajos de YARN se ralentizan de forma intermitente durante la reducción vertical, y los registros de YARN Resource Manager indican que la mayoría de sus nodos estaban en la lista de denegación durante ese tiempo, puede ajustar el límite de tiempo de espera para la retirada.

  Reduzca el valor de `spark.blacklist.decommissioning.timeout` de una hora a un minuto para que el nodo esté disponible para que otros contenedores pendientes puedan continuar con el procesamiento de las tareas.

  También debe establecer `YARN.resourcemanager.nodemanager-graceful-decommission-timeout-secs` en un valor mayor para garantizar que Amazon EMR no fuerce la terminación del nodo mientras la “Tarea de Spark” más larga siga ejecutándose en el nodo. El valor predeterminado actual son 60 minutos, lo que significa que YARN fuerza la terminación del contenedor transcurridos 60 minutos una vez que el nodo entra en el estado de retirada.

  En el siguiente ejemplo de la línea de registro de YARN Resource Manager, se muestran los nodos agregado al estado de retirada:

```
2021-10-20 15:55:26,994 INFO
 org.apache.hadoop.YARN.server.resourcemanager.DefaultAMSProcessor
 (IPC Server handler 37 on default port 8030): blacklist are updated in
 Scheduler.blacklistAdditions: [ip-10-10-27-207.us-west-2.compute.internal,
 ip-10-10-29-216.us-west-2.compute.internal, ip-10-10-31-13.us-
west-2.compute.internal, ... , ip-10-10-30-77.us-west-2.compute.internal],
 blacklistRemovals: []
```

Consulte más información sobre cómo Amazon EMR se integra con la lista de denegación de YARN durante la retirada de nodos, los casos en los que los nodos de Amazon EMR se pueden agregar a la lista de denegación y cómo configurar el comportamiento de retirada de nodos de Spark.

- En el caso de las cargas de trabajo de Spark, si se desactiva el asignador dinámico de recursos (DRA) de Spark cambiando la propiedad de Spark spark.dynamicalLocation.Enabled por, se `FALSE` pueden producir problemas de escalado gestionado, por lo que los clústeres se pueden ampliar más de lo necesario para tus cargas de trabajo (hasta el máximo de cómputo). Si utilizas Managed Scaling para estas cargas de trabajo, te recomendamos que mantengas Spark DRA activado, que es el estado predeterminado de esta propiedad.
- La sobreutilización de los volúmenes de EBS puede provocar problemas de Escalado administrado. Le recomendamos que mantenga el volumen de EBS por debajo del 90 % de utilización. Para obtener más información, consulte Comportamiento y opciones de almacenamiento de instancias en Amazon EMR.
- CloudWatch Las métricas de Amazon son fundamentales para que funcione el escalado gestionado de Amazon EMR. Te recomendamos que supervises de cerca CloudWatch las estadísticas de Amazon para asegurarte de que no falten datos. Para obtener más información sobre cómo configurar CloudWatch las alarmas para detectar las métricas faltantes, consulta Uso de CloudWatch las alarmas de Amazon.
- Las operaciones de escalado administrado en los clústeres 5.30.0 y 5.30.1 sin Presto instalado pueden provocar errores en las aplicaciones o provocar que un grupo de instancias o una flota de instancias uniformes permanezcan en estado `ARRESTED`, especialmente cuando una operación de reducción vertical va seguida inmediatamente de una operación de escalado vertical.

  Como solución alternativa, elija Presto como aplicación para instalar cuando cree un clúster con las versiones 5.30.0 y 5.30.1 de Amazon EMR, incluso si su trabajo no requiere Presto.
- Al establecer el nodo principal máximo y el límite bajo demanda de Escalado administrado de Amazon EMR, tenga en cuenta las diferencias entre los grupos de instancias y las flotas de

instancias. Cada grupo de instancias se compone del mismo tipo de instancia y las mismas opciones de compra para las instancias: bajo demanda o de spot. Para cada flota de instancias, puede especificar hasta cinco tipos de instancia que se pueden aprovisionar como instancias bajo demanda e instancias de spot. Para obtener más información, consulte Crear un clúster con las flotas de instancias o grupos de instancias uniformes, Opciones de las flotas de instancias y Escenarios de asignación de nodos.

- Con Amazon EMR 5.30.0 y versiones posteriores, si quita la regla de salida predeterminada Permitir todo a 0.0.0.0/ para el grupo de seguridad principal, debe agregar una regla que permita la conectividad TCP de salida a su grupo de seguridad para el acceso al servicio en el puerto 9443. El grupo de seguridad para el acceso al servicio también debe permitir el tráfico TCP entrante en el puerto 9443 desde el grupo de seguridad principal. Para más información sobre la configuración de grupos de seguridad, consulte Grupo de seguridad administrado por Amazon EMR para la instancia principal (subredes privadas).
- Puede utilizarlo AWS CloudFormation para configurar el escalado gestionado de Amazon EMR. Para obtener más información, consulte AWS::EMR::Cluster en la Guía del usuario de AWS CloudFormation .
- Si utiliza nodos Spot, considere la posibilidad de utilizar etiquetas de nodo para evitar que Amazon EMR elimine los procesos de aplicación cuando Amazon EMR elimine los nodos Spot. Para obtener más información sobre las etiquetas de nodos, consulte Nodos de tareas.
- No se admite el etiquetado de nodos de forma predeterminada en Amazon EMR, versión 6.15 o versiones anteriores. Para más información, consulte Descripción de los nodos principales, básicos y de tareas.
- Si utiliza Amazon EMR, versión 6.15 o versiones anteriores, solo puede asignar etiquetas de nodo por tipo de nodo, como nodos básicos y nodos de tareas. Sin embargo, si utiliza Amazon EMR versión 7.0 o posteriores, puede configurar etiquetas de nodos por tipo de nodo y tipo de mercado, como “bajo demanda” y “Spot”.
- Si la demanda de procesos de aplicación aumenta y la demanda de ejecutores disminuye al restringir el proceso de aplicación a los nodos básicos, puede volver a añadir los nodos básicos y eliminar los nodos de tareas en la misma operación de cambio de tamaño. Para obtener más información, consulte Descripción de la estrategia y los escenarios de asignación de nodos.
- Amazon EMR no etiqueta los nodos de tareas, por lo que no puede configurar las propiedades de YARN para restringir los procesos de la aplicación únicamente a los nodos de tareas. Sin embargo, si desea utilizar los tipos de mercado como etiquetas de nodos, puede utilizar las etiquetas `SPOT` o `ON_DEMAND` para ubicar los procesos de la aplicación. No recomendamos utilizar nodos de Spot para los procesos principales de la aplicación.

- Al utilizar etiquetas de nodo, el total de unidades en ejecución del clúster puede superar temporalmente el máximo del conjunto del cómputo establecido en su política de escalado administrado, mientras que Amazon EMR retira algunas de sus instancias. El total de unidades solicitadas siempre se mantendrá igual o inferior al cómputo máximo de su política.
- El escalado administrado solo admite las etiquetas de nodo `ON_DEMAND` y `SPOT` o `CORE` y `TASK`. No se admiten las etiquetas de nodo personalizadas.
- Amazon EMR crea etiquetas de nodo al crear el clúster y aprovisionar los recursos. Amazon EMR no admite la adición de etiquetas de nodo al volver a configurar el clúster. Tampoco puede modificar las etiquetas de los nodos al configurar el escalado administrado después de lanzar el clúster.
- El escalado administrado escala los nodos básicos y de tareas de forma independiente en función del proceso de aplicación y de la demanda del ejecutor. Para evitar problemas de pérdida de datos en el HDFS durante la reducción vertical básica, siga la práctica estándar para los nodos básicos. Para obtener más información sobre las prácticas recomendadas sobre los nodos básicos y la replicación de HDFS, consulte Consideraciones y prácticas recomendadas.
- No puede colocar tanto el proceso de la aplicación como los ejecutores únicamente en el nodo `core` o `ON_DEMAND`. Si quiere añadir tanto el proceso de la aplicación como los ejecutores en uno de los nodos, no utilice la configuración `yarn.node-labels.am.default-node-label-expression`.

  Por ejemplo, para colocar el proceso de la aplicación y los ejecutores en nodos `ON_DEMAND`, defina el cómputo máximo en el mismo valor que el máximo del nodo `ON_DEMAND`. Elimine también la configuración `yarn.node-labels.am.default-node-label-expression`.

  Para añadir tanto el proceso de la aplicación como los ejecutores en los nodos `core`, elimine la configuración `yarn.node-labels.am.default-node-label-expression`.
- Cuando utilice el escalado administrado con etiquetas de nodo, establezca la propiedad `yarn.scheduler.capacity.maximum-am-resource-percent: 1` si planea ejecutar varias aplicaciones en paralelo. De este modo, se asegura de que los procesos de su aplicación utilicen al máximo los nodos `ON_DEMAND` o `CORE` disponibles.
- Si utiliza el escalado administrado con etiquetas de nodos, establezca la propiedad `yarn.resourcemanager.decommissioning.timeout` en un valor que sea más largo que el de la aplicación que se ejecute durante más tiempo en el clúster. De este modo, se reduce la posibilidad de que el escalado administrado por Amazon EMR necesite reprogramar las aplicaciones para que vuelvan a ponerse en marcha los nodos `CORE` o `ON_DEMAND`.

- Para reducir el riesgo de que se produzcan errores en las aplicaciones debido a la pérdida de datos de mezcla, Amazon EMR recopila métricas del clúster para determinar los nodos que tienen datos de mezcla transitoria existentes de la etapa actual y anterior. En raras ocasiones, las métricas pueden seguir reportando datos obsoletos de aplicaciones que ya se han completado o cancelado. Esto puede afectar a la reducción puntual de las instancias del clúster. Para los clústeres que tienen una gran cantidad de datos de mezcla, considere la posibilidad de utilizar las versiones 6.13 y posteriores de EMR.

Historial de características

En esta tabla se enumeran las actualizaciones de la capacidad de Escalado administrado de Amazon EMR.

Fecha de publicación	Funcionalidad	Versiones de Amazon EMR
20 de noviembre de 2024	El escalado gestionado está disponible en las regiones de `il-central-1` Israel (Tel Aviv), `me-central-1` Oriente Medio (Emiratos Árabes Unidos) y `ap-northeast-3` Asia Pacífico (Osaka).	5.30.0 y 6.1.0 y versiones posteriores
15 de noviembre de 2024	El escalado gestionado está disponible en la región de `eu-central-2` Europa (Zúrich).	5.30.0 y 6.1.0 y versiones posteriores
20 de agosto de 2024	Las etiquetas de nodos ahora están disponibles en el escalado administrado, por lo que puede etiquetar las instancias según el tipo de mercado o el tipo de nodo para mejorar el escalado automático.	7.2.0 y versiones posteriores

Fecha de publicación	Funcionalidad	Versiones de Amazon EMR
31 de marzo de 2024	El escalado administrado está disponible en la región `ap-south-2` : Asia-Pacífico (Hyderabad).	6.14.0 y versiones posteriores
13 de febrero de 2024	El escalado administrado está disponible en la región `eu-south-2` : Europa (España).	6.14.0 y versiones posteriores
10 de octubre de 2023	El Escalado administrado está disponible en la región de Asia-Pacífico (Yakarta) `ap-southeast-3` .	6.14.0 y versiones posteriores
28 de julio de 2023	Mejora del escalado administrado para cambiar a un grupo de instancias de tareas diferente al escalar verticalmente cuando Amazon EMR experimenta un retraso en el escalado vertical con el grupo de instancias actual.	5.34.0 y versiones posteriores, 6.4.0 y versiones posteriores
16 de junio de 2023	Mejora del escalado administrado para detectar los nodos que ejecutan la aplicación maestra, de modo que esos nodos no se reduzcan verticalmente. Para obtener más información, consulte Descripción de la estrategia y los escenarios de asignación de nodos de Amazon EMR.	5.34.0 y versiones posteriores, 6.4.0 y versiones posteriores

Fecha de publicación	Funcionalidad	Versiones de Amazon EMR
21 de marzo de 2022	Se agregó el reconocimiento de datos de mezclas aleatoria s de Spark que se usa al reducir verticalmente los clústeres. En el caso de los clústeres de Amazon EMR con Apache Spark y la característica de escalado administrado habilitada, Amazon EMR supervisa de forma continua los ejecutore s de Spark y las ubicacion es de datos de mezclas aleatorias intermedias. Con esta información, Amazon EMR solo reduce verticalm ente las instancias infrautil izadas que no contienen datos de mezclas aleatoria s utilizados activamente. Esto evita el recálculo de los datos de mezclas aleatoria s perdidos, lo que ayuda a reducir los costos y a mejorar el rendimiento laboral. Para más información, consulte Spark Programming Guide.	5.34.0 y versiones posteriores, 6.4.0 y versiones posteriores

Configuración del escalado administrado para Amazon EMR

En las siguientes secciones se explica cómo lanzar un clúster de EMR que utilice el escalado gestionado con AWS Management Console AWS SDK para Java, el o el. AWS Command Line Interface

Temas

- Utilice el AWS Management Console para configurar el escalado gestionado
- Utilícela AWS CLI para configurar el escalado gestionado
- Se utiliza AWS SDK para Java para configurar el escalado gestionado

Utilice el AWS Management Console para configurar el escalado gestionado

Puede usar la consola de Amazon EMR para configurar el escalado administrado al crear un clúster o para cambiar una política de escalado administrado para un clúster en ejecución.

Console

Para configurar el escalado administrado al crear un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. Seleccione la versión emr-5.30.0 o posterior de Amazon EMR, excepto la versión emr-6.0.0.
4. En Opción de escalado y aprovisionamiento del clúster, seleccione Usar escalado administrado de EMR. Especifique la cantidad mínima y máxima de instancias, la cantidad máxima de instancias de nodos principales y la cantidad máxima de instancias bajo demanda.
5. Elija cualquier otra opción que se aplique a su clúster.
6. Para lanzar el clúster, elija Crear clúster.

Para configurar el escalado administrado en un clúster existente con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar.
3. En la pestaña Instancias de la página de detalles del clúster, busque la sección Configuración del grupo de instancias. Seleccione Editar el escalado del clúster para especificar nuevos valores para la cantidad mínima y máxima de instancias y el límite bajo demanda.

Utilícela AWS CLI para configurar el escalado gestionado

Puede utilizar AWS CLI los comandos de Amazon EMR para configurar el escalado gestionado al crear un clúster. Puede utilizar una sintaxis abreviada, especificando la configuración JSON insertada dentro de los comandos pertinentes o puede hacer referencia a un archivo que contenga la configuración JSON. También puede aplicar una política de escalado administrado a un clúster existente y eliminar una política de escalado administrado que se haya aplicado anteriormente. Además, puede recuperar detalles de una configuración de política de escalado desde un clúster en ejecución.

Habilitar el escalado administrado durante el lanzamiento del clúster

Puede habilitar el escalado administrado durante el lanzamiento del clúster, como se muestra en el siguiente ejemplo.

```
aws emr create-cluster \
 --service-role EMR_DefaultRole \
 --release-label emr-7.10.0 \
 --name EMR_Managed_Scaling_Enabled_Cluster \
 --applications Name=Spark Name=Hbase \
 --ec2-attributes KeyName=keyName,InstanceProfile=EMR_EC2_DefaultRole \
 --instance-groups InstanceType=m4.xlarge,InstanceGroupType=MASTER,InstanceCount=1
 InstanceType=m4.xlarge,InstanceGroupType=CORE,InstanceCount=2 \
 --region us-east-1 \
 --managed-scaling-policy
 ComputeLimits='{MinimumCapacityUnits=2,MaximumCapacityUnits=4,UnitType=Instances}'
```

También puede especificar una configuración de política administrada mediante la managed-scaling-policy opción -- cuando utilice`create-cluster`.

Aplicación de una política de escalado administrado a un clúster existente

Puede aplicar una política de escalado administrado a un clúster existente, como se muestra en el siguiente ejemplo.

```
aws emr put-managed-scaling-policy
--cluster-id j-123456
--managed-scaling-policy ComputeLimits='{MinimumCapacityUnits=1,
MaximumCapacityUnits=10,  MaximumOnDemandCapacityUnits=10, UnitType=Instances}'
```

También puede aplicar una política de escalado administrado a un clúster existente mediante el comando `aws emr put-managed-scaling-policy`. En el siguiente ejemplo se utiliza una referencia a un archivo JSON, `managedscaleconfig.json`, que especifica la configuración de política de escalado administrado.

```
aws emr put-managed-scaling-policy --cluster-id j-123456 --managed-scaling-policy
 file://./managedscaleconfig.json
```

En el ejemplo siguiente se muestra el contenido del archivo `managedscaleconfig.json`, que define la política de escalado administrado.

```
{
    "ComputeLimits": {
        "UnitType": "Instances",
        "MinimumCapacityUnits": 1,
        "MaximumCapacityUnits": 10,
        "MaximumOnDemandCapacityUnits": 10
    }
}
```

Recuperación de una configuración de política de escalado administrado

El comando `GetManagedScalingPolicy` recupera la configuración de la política. Por ejemplo, el comando siguiente recupera la configuración para el clúster con un ID de clúster de `j-123456`.

```
aws emr get-managed-scaling-policy --cluster-id j-123456
```

El comando produce el siguiente resultado de ejemplo.

```
{
   "ManagedScalingPolicy": {
      "ComputeLimits": {
         "MinimumCapacityUnits": 1,
         "MaximumOnDemandCapacityUnits": 10,
         "MaximumCapacityUnits": 10,
         "UnitType": "Instances"
      }
   }
}
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Eliminación de una política de escalado administrado

El comando `RemoveManagedScalingPolicy` elimina la configuración de política. Por ejemplo, el comando siguiente recupera la configuración del clúster con un ID de clúster de `j-123456`.

```
aws emr remove-managed-scaling-policy --cluster-id j-123456
```

Se utiliza AWS SDK para Java para configurar el escalado gestionado

En el siguiente fragmento de programa se muestra cómo configurar el escalado administrado mediante AWS SDK para Java:

```
package com.amazonaws.emr.sample;

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.regions.Regions;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduce;
import com.amazonaws.services.elasticmapreduce.AmazonElasticMapReduceClientBuilder;
import com.amazonaws.services.elasticmapreduce.model.Application;
import com.amazonaws.services.elasticmapreduce.model.ComputeLimits;
import com.amazonaws.services.elasticmapreduce.model.ComputeLimitsUnitType;
import com.amazonaws.services.elasticmapreduce.model.InstanceGroupConfig;
import com.amazonaws.services.elasticmapreduce.model.JobFlowInstancesConfig;
import com.amazonaws.services.elasticmapreduce.model.ManagedScalingPolicy;
import com.amazonaws.services.elasticmapreduce.model.RunJobFlowRequest;
import com.amazonaws.services.elasticmapreduce.model.RunJobFlowResult;

public class CreateClusterWithManagedScalingWithIG {

 public static void main(String[] args) {
  AWSCredentials credentialsFromProfile = getCreadentials("AWS-Profile-Name-Here");

  /**
```

```
  * Create an Amazon EMR client with the credentials and region specified in order to
create the cluster
  */
 AmazonElasticMapReduce emr = AmazonElasticMapReduceClientBuilder.standard()
  .withCredentials(new AWSStaticCredentialsProvider(credentialsFromProfile))
  .withRegion(Regions.US_EAST_1)
  .build();

 /**
  * Create Instance Groups - Primary, Core, Task
  */
 InstanceGroupConfig instanceGroupConfigMaster = new InstanceGroupConfig()
   .withInstanceCount(1)
   .withInstanceRole("MASTER")
   .withInstanceType("m4.large")
   .withMarket("ON_DEMAND");

 InstanceGroupConfig instanceGroupConfigCore = new InstanceGroupConfig()
  .withInstanceCount(4)
  .withInstanceRole("CORE")
  .withInstanceType("m4.large")
  .withMarket("ON_DEMAND");

 InstanceGroupConfig instanceGroupConfigTask = new InstanceGroupConfig()
  .withInstanceCount(5)
  .withInstanceRole("TASK")
  .withInstanceType("m4.large")
  .withMarket("ON_DEMAND");

 List<InstanceGroupConfig> igConfigs = new ArrayList<>();
 igConfigs.add(instanceGroupConfigMaster);
 igConfigs.add(instanceGroupConfigCore);
 igConfigs.add(instanceGroupConfigTask);

       /**
        *  specify applications to be installed and configured when Amazon EMR creates
the cluster
        */
 Application hive = new Application().withName("Hive");
 Application spark = new Application().withName("Spark");
 Application ganglia = new Application().withName("Ganglia");
 Application zeppelin = new Application().withName("Zeppelin");

 /**
```

```
  * Managed Scaling Configuration -
        * Using UnitType=Instances for clusters composed of instance groups
  *
        * Other options are:
        * UnitType = VCPU ( for clusters composed of instance groups)
        * UnitType = InstanceFleetUnits ( for clusters composed of instance fleets)
        **/
 ComputeLimits computeLimits = new ComputeLimits()
   .withMinimumCapacityUnits(1)
   .withMaximumCapacityUnits(20)
   .withUnitType(ComputeLimitsUnitType.Instances);

 ManagedScalingPolicy managedScalingPolicy = new ManagedScalingPolicy();
 managedScalingPolicy.setComputeLimits(computeLimits);

 // create the cluster with a managed scaling policy
 RunJobFlowRequest request = new RunJobFlowRequest()
         .withName("EMR_Managed_Scaling_TestCluster")
         .withReleaseLabel("emr-7.10.0")          // Specifies the version label for
the Amazon EMR release; we recommend the latest release
         .withApplications(hive,spark,ganglia,zeppelin)
         .withLogUri("s3://path/to/my/emr/logs")  // A URI in S3 for log files is
required when debugging is enabled.
         .withServiceRole("EMR_DefaultRole")      // If you use a custom IAM service
role, replace the default role with the custom role.
         .withJobFlowRole("EMR_EC2_DefaultRole")  // If you use a custom Amazon EMR
role for EC2 instance profile, replace the default role with the custom Amazon EMR
role.
         .withInstances(new JobFlowInstancesConfig().withInstanceGroups(igConfigs)
             .withEc2SubnetId("subnet-123456789012345")
             .withEc2KeyName("my-ec2-key-name")
             .withKeepJobFlowAliveWhenNoSteps(true))
         .withManagedScalingPolicy(managedScalingPolicy);
   RunJobFlowResult result = emr.runJobFlow(request);

   System.out.println("The cluster ID is " + result.toString());
}

public static AWSCredentials getCredentials(String profileName) {
 // specifies any named profile in .aws/credentials as the credentials provider
 try {
  return new ProfileCredentialsProvider("AWS-Profile-Name-Here")
    .getCredentials();
       } catch (Exception e) {
```

```
            throw new AmazonClientException(
                    "Cannot load credentials from .aws/credentials file. " +
                    "Make sure that the credentials file exists and that the profile
 name is defined within it.",
                    e);
        }
 }

 public CreateClusterWithManagedScalingWithIG() { }
}
```

Escalado avanzado para Amazon EMR

A partir de Amazon EMR en la EC2 versión 7.0, puede aprovechar Advanced Scaling para controlar la utilización de los recursos del clúster. El escalado avanzado introduce una escala de utilización-rendimiento para ajustar la utilización de los recursos y el nivel de rendimiento de acuerdo con las necesidades de su empresa. El valor que establezca determinará si el clúster se centra más en la conservación de los recursos o en la ampliación para gestionar cargas de trabajo sensibles service-level-agreement (SLA), en las que es fundamental completarlo rápidamente. Cuando se ajusta el valor de escalado, el escalado gestionado interpreta su intención y escala de forma inteligente para optimizar los recursos. Para obtener más información sobre el escalado gestionado, consulte Configurar el escalado gestionado para Amazon EMR.

Configuración de escalado avanzada

El valor que ha establecido para Advanced Scaling optimiza el clúster en función de sus necesidades. Los valores oscilan entre 1 y 100. Los valores posibles son 1, 25, 50, 75 y 100. Si establece el índice en valores distintos de estos, se producirá un error de validación.

Los valores de escalado se corresponden con las estrategias de utilización de los recursos. En la siguiente lista se definen varias de ellas:

- Utilización optimizada [1]: esta configuración evita el sobreaprovisionamiento de recursos. Utilice un valor bajo cuando desee mantener los costos bajos y priorizar la utilización eficiente de los recursos. Hace que el clúster se amplíe de forma menos agresiva. Esto funciona bien cuando se producen picos de carga de trabajo frecuentes y no se desea que los recursos aumenten demasiado rápido.
- Equilibrado [50]: equilibra la utilización de los recursos y el rendimiento laboral. Esta configuración es adecuada para cargas de trabajo estables en las que la mayoría de las etapas tienen un tiempo de ejecución estable. También es adecuado para cargas de trabajo con una combinación de

etapas de corta y larga duración. Te recomendamos que empieces con esta configuración si no estás seguro de cuál elegir.

- Rendimiento optimizado [100]: esta estrategia prioriza el rendimiento. El clúster se amplía agresivamente para garantizar que los trabajos se completen rápidamente y cumplan los objetivos de rendimiento. El rendimiento optimizado es adecuado para cargas de trabajo sensibles service-level-agreement (SLA) en las que un tiempo de ejecución rápido es fundamental.

> **Note**
>
> Los valores intermedios disponibles proporcionan un punto medio entre las estrategias para ajustar con precisión el comportamiento de escalado avanzado del clúster.

Ventajas del escalado avanzado

Dado que su entorno y sus requisitos varían, como los cambios en los volúmenes de datos, los ajustes de los objetivos de costes y las implementaciones de los acuerdos de nivel de servicio, el escalado de clústeres puede ayudarle a ajustar la configuración del clúster para alcanzar sus objetivos. Los beneficios clave incluyen:

- Control granular mejorado: la introducción de la configuración de utilización-rendimiento le permite ajustar fácilmente el comportamiento de escalado del clúster de acuerdo con sus requisitos. Puede ampliarlo para satisfacer la demanda de recursos informáticos o reducirlo para ahorrar recursos, en función de sus patrones de uso.
- Optimización de costes mejorada: puede elegir un valor de utilización bajo según lo exijan los requisitos para cumplir más fácilmente sus objetivos de costes.

Cómo empezar con la optimización

Instalación y configuración

Siga estos pasos para establecer el índice de rendimiento y optimizar su estrategia de escalado.

1. El siguiente comando actualiza un clúster existente con la estrategia de `[1]` escalado optimizada para el uso:

```
aws emr put-managed-scaling-policy --cluster-id 'cluster-id' \
```

```
 --managed-scaling-policy '{
  "ComputeLimits": {
    "UnitType": "Instances",
    "MinimumCapacityUnits": 1,
    "MaximumCapacityUnits": 2,
    "MaximumOnDemandCapacityUnits": 2,
    "MaximumCoreCapacityUnits": 2
  },
  "ScalingStrategy": "ADVANCED",
  "UtilizationPerformanceIndex": "1"
}' \
 --region "region-name"
```

Los atributos `ScalingStrategy` y `UtilizationPerformanceIndex` son nuevos y relevantes para la optimización del escalado. Puede seleccionar diferentes estrategias de escalado estableciendo los valores correspondientes (1, 25, 50, 75 y 100) para el `UtilizationPerformanceIndex` atributo en la política de escalado gestionado.

2. Para volver a la estrategia de escalado administrado predeterminada, ejecute el `put-managed-scaling-policy` comando sin incluir los atributos y. `ScalingStrategy` `UtilizationPerformanceIndex` (Esto es opcional). En este ejemplo se muestra cómo hacerlo:

```
aws emr put-managed-scaling-policy \
--cluster-id 'cluster-id' \
--managed-scaling-policy '{"ComputeLimits":
{"UnitType":"Instances","MinimumCapacityUnits":1,"MaximumCapacityUnits":2,"MaximumOnDemandC
 \
--region "region-name"
```

Uso de métricas de monitoreo para realizar un seguimiento de la utilización del clúster

A partir de la versión 7.3.0 de EMR, Amazon EMR publica cuatro métricas nuevas relacionadas con la memoria y la CPU virtual. Puede utilizarlas para medir la utilización del clúster en todas las estrategias de escalado. Estas métricas están disponibles para cualquier caso de uso, pero puede utilizar los detalles que se proporcionan aquí para supervisar el escalado avanzado.

Entre las métricas útiles disponibles se incluyen las siguientes:

- YarnContainersUsedMemoryGBSeconds— Cantidad de memoria consumida por las aplicaciones administradas por YARN.

- YarnContainersTotalMemoryGBSeconds— Capacidad de memoria total asignada a YARN dentro del clúster.
- YarnNodesUsedVCPUSeconds— Número total de segundos de CPU virtual para cada aplicación gestionada por YARN.
- YarnNodesTotalVCPUSeconds— El total de segundos de CPU virtual acumulados para la memoria consumida, incluido el intervalo de tiempo en el que Yarn no está listo.

Puede analizar las métricas de los recursos mediante Amazon CloudWatch Logs Insights. Las funciones incluyen un lenguaje de consultas especialmente diseñado que le ayuda a extraer métricas específicas para el uso y el escalamiento de los recursos.

La siguiente consulta, que se puede ejecutar en la Amazon CloudWatch consola, utiliza la matemática métrica para calcular la utilización media de la memoria (e1) dividiendo la suma acumulada de la memoria consumida (e2) por la suma acumulada de la memoria total (e3):

```
{
    "metrics": [
        [ { "expression": "e2/e3", "label": "Average Mem Utilization", "id": "e1",
 "yAxis": "right" } ],
        [ { "expression": "RUNNING_SUM(m1)", "label": "RunningTotal-
YarnContainersUsedMemoryGBSeconds", "id": "e2", "visible": false } ],
        [ { "expression": "RUNNING_SUM(m2)", "label": "RunningTotal-
YarnContainersTotalMemoryGBSeconds", "id": "e3", "visible": false } ],
        [ "AWS_EMR_ManagedResize", "YarnContainersUsedMemoryGBSeconds", "ACCOUNT_ID",
 "793684541905", "COMPONENT", "ManagerService", "JOB_FLOW_ID", "cluster-id", { "id":
 "m1", "label": "YarnContainersUsedMemoryGBSeconds" } ],
        [ ".", "YarnContainersTotalMemoryGBSeconds", ".", ".", ".", ".", ".", ".",
 { "id": "m2", "label": "YarnContainersTotalMemoryGBSeconds" } ]
    ],
    "view": "timeSeries",
    "stacked": false,
    "region": "region",
    "period": 60,
    "stat": "Sum",
    "title": "Memory Utilization"
}
```

Para consultar los registros, puede seleccionarlos CloudWatch en la AWS consola. Para obtener más información sobre cómo escribir consultas CloudWatch, consulte Análisis de datos de registro con CloudWatch Logs Insights en la Guía del usuario de Amazon CloudWatch Logs.

La siguiente imagen muestra estas métricas para un clúster de muestra:

Consideraciones y limitaciones

- La eficacia de las estrategias de escalado puede variar en función de las características únicas de la carga de trabajo y de la configuración del clúster. Le recomendamos que experimente con la configuración de escalado para determinar un valor de índice óptimo para su caso de uso.
- Amazon EMR Advanced Scaling es especialmente adecuado para cargas de trabajo por lotes. Para cargas de trabajo de SQL/almacenamiento de datos y streaming, recomendamos utilizar la estrategia de escalado administrado predeterminada para obtener un rendimiento óptimo.
- La estrategia de escalado optimizada para el rendimiento permite una ejecución más rápida de los trabajos al mantener una gran cantidad de recursos de cómputo durante un período más prolongado que la estrategia de escalado administrado predeterminada. Este modo prioriza el escalado rápido para satisfacer la demanda de recursos, lo que se traduce en una finalización más rápida del trabajo. Esto podría generar costos más altos en comparación con la estrategia predeterminada.
- En los casos en que el clúster ya esté optimizado y se utilice en su totalidad, es posible que habilitar el escalado avanzado no proporcione beneficios adicionales. En algunas situaciones, habilitar el escalado avanzado puede generar un aumento de los costos, ya que las cargas de trabajo pueden durar más. En estos casos, recomendamos utilizar la estrategia de escalado gestionado predeterminada para garantizar una asignación de recursos y una rentabilidad óptimas.
- En el contexto del escalado gestionado, se hace más hincapié en la utilización de los recursos y no en el tiempo de ejecución, ya que la configuración pasa de estar optimizada para el rendimiento [100] a optimizada para la utilización [1]. Sin embargo, es importante tener en cuenta que los resultados pueden variar según la naturaleza de la carga de trabajo y la topología del clúster. Para garantizar resultados óptimos para su caso de uso, le recomendamos encarecidamente que pruebe las estrategias de escalado con sus cargas de trabajo para determinar la configuración más adecuada.
- The solo PerformanceUtilizationIndexacepta los siguientes valores:
  - 1
  - 25
  - 50
  - 75
  - 100

Cualquier otro valor enviado provocará un error de validación.

Descripción de la estrategia y los escenarios de asignación de nodos de Amazon EMR

En esta sección se ofrece información general sobre la estrategia de asignación de nodos y los escenarios de escalado comunes que puede utilizar con Escalado administrado de Amazon EMR.

Estrategia de asignación de nodos

Escalado administrado de Amazon EMR asigna los nodos principales y de tarea en función de las siguientes estrategias de escalado y reducción verticales:

Estrategia de escalado vertical

- Para las versiones 7.2 y posteriores de Amazon EMR, el escalado administrado añade primero los nodos en función de las etiquetas de los nodos y de la propiedad YARN, que restringe el proceso de aplicación.
- Por ejemplo, si utiliza la versión 7.2 o posteriores de Amazon EMR, si habilita las etiquetas de nodos y restringe el proceso de aplicación a los nodos `CORE`, el escalado administrado de Amazon EMR escala verticalmente los nodos básicos y los nodos de tarea si aumenta la demanda del proceso de aplicación y aumenta la demanda del ejecutor. Igualmente, si habilitó las etiquetas de nodos y restringió los procesos de aplicación a los nodos `ON_DEMAND`, el escalado administrado escala verticalmente los nodos bajo demanda si aumenta la demanda del proceso de aplicación y se escalan verticalmente los nodos Spot si aumenta la demanda del ejecutor.
- Si las etiquetas de nodos no están habilitadas, la ubicación de los procesos de aplicación no está restringida a ningún nodo o tipo de mercado.
- Al usar etiquetas de nodos, el escalado administrado puede escalar verticalmente y reducir verticalmente diferentes grupos de instancias y flotas de instancias en la misma operación de cambio de tamaño. Por ejemplo, en un escenario en el que `instance_group1` tiene un nodo `ON_DEMAND` y `instance_group2` tiene un nodo `SPOT`, y las etiquetas de nodo están habilitadas y los procesos de aplicación están restringidos a los nodos con la etiqueta `ON_DEMAND`. El escalado administrado reducirá verticalmente `instance_group1` y escalará verticalmente `instance_group2` si la demanda del proceso de aplicación disminuye y la demanda del ejecutor aumenta.
- Cuando Amazon EMR experimenta un retraso en el escalado vertical con el grupo de instancias actual, los clústeres que usan el escalado administrado cambian automáticamente a un grupo de instancias de tareas diferente.

- Si el parámetro `MaximumCoreCapacityUnits` está establecido, Amazon EMR escala los nodos principales hasta que las unidades principales alcanzan el límite máximo permitido. Toda la capacidad restante se agrega a los nodos de tarea.
- Si el parámetro `MaximumOnDemandCapacityUnits` está establecido, Amazon EMR escala el clúster mediante las instancias bajo demanda hasta que las unidades bajo demanda alcanzan el límite máximo permitido. Toda la capacidad restante se agrega mediante instancias de spot.
- Si se establecen los parámetros `MaximumCoreCapacityUnits` y `MaximumOnDemandCapacityUnits`, Amazon EMR tiene en cuenta ambos límites durante el escalado.

  Por ejemplo, si `MaximumCoreCapacityUnits` es inferior a `MaximumOnDemandCapacityUnits`, Amazon EMR primero escala los nodos principales hasta alcanzar el límite de capacidad principal. Para la capacidad restante, Amazon EMR utiliza primero las instancias bajo demanda para escalar los nodos de tarea hasta alcanzar el límite bajo demanda y, a continuación, utiliza las instancias de spot para los nodos de tarea.

Estrategia de reducción vertical

- De forma similar a la estrategia de escalado vertical, Amazon EMR elimina los nodos en función de las etiquetas de los nodos. Para más información sobre las etiquetas de los nodos, consulte Descripción de los tipos de nodos: principales, básicos y de tareas.
- Si no ha habilitado las etiquetas de nodo, el escalado administrado elimina los nodos de tarea y, a continuación, elimina los nodos básicos hasta que alcanza la capacidad objetivo de reducción vertical deseada. El escalado administrado nunca reduce verticalmente el clúster por debajo de las restricciones mínimas de la política de escalado administrado.
- Las versiones 5.34.0 y posteriores de Amazon EMR, y las versiones 6.4.0 y posteriores de Amazon EMR, admiten el reconocimiento de datos aleatorio de Spark, lo que evita que una instancia se reduzca mientras Managed Scaling conoce los datos aleatorios existentes. Para más información sobre las operaciones de mezclas aleatorias, consulte Spark Programming Guide. Managed Scaling hace todo lo posible para evitar reducir la escala de los nodos con datos mezclados de la fase actual y anterior de cualquier aplicación de Spark activa, durante un máximo de 30 minutos. Esto ayuda a minimizar la pérdida de datos aleatoria involuntaria, evitando la necesidad de volver a intentar el trabajo y tener que volver a calcular los datos intermedios. Sin embargo, no se garantiza la prevención de la pérdida de datos de forma aleatoria. Para garantizar una protección garantizada, recomendamos utilizar el código aleatorio en los clústeres con la

etiqueta de publicación 7.4 o superior. Consulta a continuación cómo configurar la protección aleatoria garantizada.

- El escalado administrado elimina primero los nodos de tarea y, a continuación, elimina los nodos básicos hasta que se alcanza la capacidad objetivo de reducción vertical deseada. El clúster nunca se escala por debajo de las restricciones mínimas de la política de escalado administrado.
- En el caso de los clústeres que se lanzan con las versiones 5.x, 5.34.0 y posteriores, y las versiones 6.x, 6.4.0 y posteriores de Amazon EMR, el Escalado administrado de Amazon EMR no reduce verticalmente los nodos que tienen `ApplicationMaster` para Apache Spark en ejecución. Esto minimiza los errores y los reintentos en los trabajos, lo que ayuda a mejorar el rendimiento de los trabajos y a reducir los costos. Para confirmar qué nodos del clúster se ejecutan `ApplicationMaster`, visite el servidor del historial de Spark y filtre el controlador en la pestaña Ejecutores del ID de aplicación de Spark.
- Si bien el escalado inteligente con EMR Managed Scaling minimiza la pérdida de datos aleatorios para Spark, puede haber casos en los que los datos aleatorios transitorios no estén protegidos durante una reducción de escala. Para mejorar la resiliencia de los datos aleatorizados durante la reducción de escala, recomendamos activar Graceful Decommissioning for Shuffle Data en YARN. Cuando la función Graceful Decommissioning for Shuffle Data está activada en YARN, los nodos seleccionados para la reducción de escala que contengan datos aleatorizados pasarán al estado de desmantelamiento y seguirán distribuyendo archivos aleatorios. El YARN ResourceManager espera a que los nodos notifiquen que no hay ningún archivo aleatorio presente antes de eliminarlos del clúster.
  - La versión 6.11.0 y las versiones posteriores de Amazon EMR admiten el desmantelamiento gradual basado en Yarn para los datos de Hive Shuffle, tanto para los controladores Tez como para los Shuffle. MapReduce
    - Active `yarn.resourcemanager.decommissioning-nodes-watcher.wait-for-shuffle-data` Graceful `true` Decommissioning for Shuffle Data configurando en.
  - La versión 7.4.0 y las versiones posteriores de Amazon EMR admiten el desmantelamiento gradual basado en YARN para los datos de reproducción aleatoria de Spark cuando el servicio de reproducción aleatoria externo está habilitado (activado de forma predeterminada cuando la EMR está activada). EC2
    - El comportamiento predeterminado del servicio aleatorio externo de Spark, cuando se ejecuta Spark en Yarn, es que Yarn elimine los archivos de reproducción aleatoria de aplicaciones NodeManager al finalizar la aplicación. Esto puede afectar a la velocidad de desmantelamiento de los nodos y a la utilización del cómputo. En el caso de las aplicaciones de ejecución prolongada, considere la posibilidad de configurar esta opción

`spark.shuffle.service.removeShuffle` `true` para eliminar los archivos aleatorios que ya no se utilizan para poder desmantelar más rápidamente los nodos sin datos de reproducción aleatoria activos.

- Si el `yarn.nodemanager.shuffledata-monitor.interval-ms` indicador o el han `spark.dynamicAllocation.executorIdleTimeout` cambiado con respecto a los valores predeterminados, asegúrese de que la condición se `spark.dynamicAllocation.executorIdleTimeout > yarn.nodemanager.shuffledata-monitor.interval-ms` mantenga `true` actualizando el indicador necesario.

Si el clúster no tiene ninguna carga, Amazon EMR cancela la adición de nuevas instancias de una evaluación anterior y realiza operaciones de reducción vertical. Si el clúster tiene una carga elevada, Amazon EMR cancela la eliminación de instancias y realiza operaciones de escalado vertical.

Consideraciones sobre la asignación de nodos

Le recomendamos que utilice la opción de compra bajo demanda para los nodos principales a fin de evitar la pérdida de datos de HDFS en caso de recuperación de spot. Puede utilizar la opción de compra puntual para los nodos de tarea a fin de reducir los costos y agilizar la ejecución de los trabajos cuando se agreguen más instancias de spot a los nodos de tarea.

Escenarios de asignación de nodos

Puede crear varios escenarios de escalado en función de sus necesidades configurando los parámetros máximo, mínimo, límite bajo demanda y máximo de nodos principales en diferentes combinaciones.

Escenario 1: escalar únicamente los nodos principales

Para escalar únicamente los nodos principales, los parámetros de escalado administrado deben cumplir los siguientes requisitos:

- El límite bajo demanda es igual al límite máximo.
- El máximo de nodos principales es igual al límite máximo.

Cuando no se especifican los parámetros del límite bajo demanda y el máximo de nodos principales, ambos parámetros se establecen de forma predeterminada en el límite máximo.

Este escenario no es aplicable si utiliza el escalado administrado con etiquetas de nodos y restringe los procesos de su aplicación para que solo se ejecuten en los nodos `CORE`, ya que el escalado administrado escala los nodos de tareas para adaptarlos a la demanda de los ejecutores.

En los siguientes ejemplos, se muestra el escenario de escalado de nodos principales únicamente.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Grupos de instancias Principal: 1 bajo demanda Tarea: 1 bajo demanda y 1 de spot	`UnitType`: instancias `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 20	Escale entre 1 y 20 instancias o unidades de flota de instancias en los nodos principales mediante el tipo bajo demanda. No se puede escalar en los nodos de tarea. Si utiliza el escalado administrado con etiquetas de nodos y restringe los procesos de las aplicaciones a nodos `ON_DEMAND`, el clúster escalará de 1 a 20 instancias o unidades de flota de instancias en los nodos `CORE` utilizand
Flotas de instancias Principal: 1 bajo demanda Tarea: 1 bajo demanda y 1 de spot	UnitType: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 20	

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
		o el tipo `Spot` o `On-Demand` , en función del tipo de demanda.

Escenario 2: escalar únicamente los nodos de tarea

Para escalar únicamente los nodos de tarea, los parámetros de escalado administrado deben cumplir el siguiente requisito:

- El máximo de nodos principales debe ser igual al límite mínimo.

En los siguientes ejemplos, se muestra el escenario de escalado de nodos de tarea únicamente.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Grupos de instancias Principal: 2 bajo demanda Tarea: 1 de spot	`UnitType`: instancias `MinimumCapacityUnits` : 2 `MaximumCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 2	Mantenga los nodos principal es estables en 2 y escale únicamente los nodos de tarea entre 0 y 18 instancias o unidades de flota de instancias. La capacidad entre los límites mínimo y máximo se agrega
Flotas de instancias Principal: 2 bajo demanda Tarea: 1 de spot	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 2 `MaximumCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 2	

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
		únicamente a los nodos de tarea. Cuando utiliza el escalado administrado con etiquetas de nodo y restringe los procesos de sus aplicaciones a nodos ON_DEMAND, el clúster mantendrá los nodos básicos estables en 2 y solo escalará los nodos de tareas entre 0 y 18 instancias o unidades de flota de instancias que usen el tipo `On-demand` o `Spot`, según el tipo de demanda.

Escenario 3: solo las instancias bajo demanda del clúster

Para tener únicamente instancias bajo demanda, el clúster y los parámetros de escalado administrado deben cumplir el siguiente requisito:

- El límite bajo demanda es igual al límite máximo.

Si no se especifica el límite bajo demanda, el valor del parámetro se establece de forma predeterminada en el límite máximo. El valor predeterminado indica que Amazon EMR escala únicamente las instancias bajo demanda.

Si el máximo de nodos principales es inferior al límite máximo, el parámetro del máximo de nodos principales se puede utilizar para dividir la asignación de capacidad entre los nodos principales y de tarea.

Para habilitar este escenario en un clúster compuesto por grupos de instancias, todos los grupos de nodos del clúster deben usar el tipo de compra bajo demanda durante la configuración inicial.

Este escenario no es aplicable si utiliza el escalado administrado con etiquetas de nodo y restringe los procesos de las aplicaciones para que solo se ejecuten en los nodos `ON_DEMAND`, ya que el escalado gestionado escala los nodos `Spot` para adaptarse a la demanda de los ejecutores.

En los siguientes ejemplos, se muestra el escenario en el que hay instancias bajo demanda en todo el clúster.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Grupos de instancias Principal: 1 bajo demanda Tarea: 1 baja demanda	`UnitType`: instancias `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 12	Escale entre 1 y 12 instancias o unidades de flota de instancias en los nodos principales mediante el tipo bajo demanda. Escale la capacidad restante mediante el tipo bajo demanda en los nodos de tarea. No se
Flotas de instancias Principal: 1 bajo demanda Tarea: 1 baja demanda	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20	

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
	`MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 12	puede escalar con instancias de spot. Si utiliza el escalado administrado con etiquetas de nodos y restringe los procesos de las aplicaciones a los nodos `CORE`, el clúster escala entre 1 y 20 instancias o unidades de flota de instancias en los nodos `CORE` o los nodos `task` que utilizan ese tipo `ON_DEMAND`, según el tipo de demanda. El escalado en los nodos básicos no superará las 12 instancias o unidades de flota de instancias.

Escenario 4: solo las instancias de spot del clúster

Para tener únicamente instancias de spot, los parámetros de escalado administrado deben cumplir el siguiente requisito:

- El límite bajo demanda está establecido en 0.

Si el máximo de nodos principales es inferior al límite máximo, el parámetro del máximo de nodos principales se puede utilizar para dividir la asignación de capacidad entre los nodos principales y de tarea.

Para habilitar este escenario en un clúster compuesto por grupos de instancias, el grupo de instancias principales debe usar la opción de compra de spot durante la configuración inicial. Si no hay ninguna instancia de spot en el grupo de instancias de las tareas, Escalado administrado de Amazon EMR crea un grupo de tareas que utiliza instancias de spot cuando es necesario.

Este escenario no es aplicable si utiliza el escalado administrado con etiquetas de nodo y restringe los procesos de sus aplicaciones para que solo se ejecuten en los nodos `ON_DEMAND`, ya que el escalado administrado escala los nodos `ON_DEMAND` para adaptarse a la demanda de los procesos de aplicación.

En los siguientes ejemplos, se muestra el escenario en el que hay instancias de spot en todo el clúster.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Grupos de instancias Principal: 1 de spot Tarea: 1 de spot	`UnitType`: instancias `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 0	Escale entre 1 y 20 instancias o unidades de flota de instancias en los nodos principales mediante el tipo de spot. No se puede escalar con el tipo bajo demanda.
Flotas de instancias Principal: 1 de spot Tarea: 1 de spot	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1	

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
	`MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 0	Cuando utiliza el escalado administrado con etiquetas de nodos y restringe los procesos de sus aplicaciones a nodos `CORE`, el clúster escala entre 1 y 20 instancias o unidades de flota de instancias en nodos `CORE` o `TASK` que utilizan Spot, según el tipo de demanda. Amazon EMR no escala mediante el tipo `ON_DEMAND` .

Escenario 5: escalar las instancias bajo demanda en los nodos principales y las instancias de spot en los nodos de tarea

Para escalar las instancias bajo demanda en los nodos principales y las instancias de spot en los nodos de tarea, los parámetros de escalado administrado deben cumplir los siguientes requisitos:

- El límite bajo demanda debe ser igual al máximo de nodos principales.
- Tanto el límite bajo demanda como el máximo de nodos principales deben ser inferiores al límite máximo.

Para habilitar este escenario en un clúster compuesto por grupos de instancias, el grupo de nodos principales debe usar la opción de compra bajo demanda.

Este escenario no es aplicable si utiliza el escalado administrado con etiquetas de nodo y restringe los procesos de su aplicación para que solo se ejecuten en nodos `ON_DEMAND` o en nodos `CORE`.

En los siguientes ejemplos, se muestra el escenario en el que se escalan las instancias bajo demanda en los nodos principales y las instancias de spot en los nodos de tarea.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Grupos de instancias Principal: 1 bajo demanda Tarea: 1 bajo demanda y 1 de spot	`UnitType`: instancias `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 7 `MaximumCoreCapacityUnits` : 7	Escale hasta 6 unidades bajo demanda en el nodo principal , ya que ya hay una unidad bajo demanda en el nodo de tarea y el límite máximo de unidades bajo demanda es de 7. A continuación, escale hasta 13 unidades de spot en los nodos de tarea.
Flotas de instancias Principal: 1 bajo demanda Tarea: 1 bajo demanda y 1 de spot	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 7 `MaximumCoreCapacityUnits` : 7	

Escenario 6: escale las instancias **`CORE`** según la demanda del proceso de la aplicación y las instancias **`TASK`** según la demanda de los ejecutores.

Este escenario no es aplicable si utiliza el escalado administrado con etiquetas de nodo y restringe los procesos de su aplicación para que solo se ejecuten en nodos `CORE`.

Para escalar los nodos `CORE` en función de la demanda del proceso de la aplicación y los nodos `TASK` en función de la demanda del ejecutor, debe establecer las siguientes configuraciones al lanzar el clúster:

- `yarn.node-labels.enabled:true`
- `yarn.node-labels.am.default-node-label-expression: 'CORE'`

Si no especifica el límite `ON_DEMAND` y los parámetros máximos de nodos `CORE`, ambos parámetros se establecen de forma predeterminada en el límite máximo.

Si el máximo de nodos `ON_DEMAND` es inferior al límite máximo, el escalado administrado usa el parámetro máximo de nodos `ON_DEMAND` para dividir la asignación de capacidad entre los nodos `ON_DEMAND` y `SPOT`. Si establece el parámetro máximo de nodo `CORE` en un valor inferior o igual al parámetro de capacidad mínima, los nodos `CORE` permanecen estáticos con la capacidad máxima básica.

En los siguientes ejemplos, se muestra el escenario en el que se escalan las instancias CORE en función de la demanda de procesos de aplicación y las instancias TASK en función de la demanda del ejecutor.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Grupos de instancias Principal: 1 bajo demanda Tarea: 1 baja demanda	`UnitType`: instancias `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 10 `MaximumCoreCapacityUnits` : 20	Escala los nodos `CORE` entre 1 y 20 nodos en función de la demanda de procesos de aplicación del clúster, utilizando el tipo de mercado bajo demanda o Spot. Escala los nodos `TASK` en función de la demanda
Flotas de instancias Principal: 1 bajo demanda	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1	

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Tarea: 1 baja demanda	`MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 10 `MaximumCoreCapacityUnits` : 20	del ejecutor y de la capacidad restante disponible después de que Amazon EMR asigne los nodos `CORE`. La suma de los nodos `TASK` y `CORE` solicitados no superará la `maximumCapacity` de 20. La suma de los nodos básicos bajo demanda y de los nodos de tareas bajo demanda solicitados no superará la `maximumOnDemandCapacity` de 10. Los nodos básicos o de tareas adicionales utilizan el tipo de mercado o de Spot.

Escenario 7: escale las instancias **ON_DEMAND** según la demanda del proceso de aplicación y las instancias **SPOT** según la demanda de los ejecutores.

Este escenario no es aplicable si utiliza el escalado administrado con etiquetas de nodo y restringe los procesos de su aplicación para que solo se ejecuten en nodos `ON_DEMAND`.

Para escalar los nodos `ON_DEMAND` en función de la demanda del proceso de la aplicación y los nodos `SPOT` en función de la demanda del ejecutor, debe establecer las siguientes configuraciones al lanzar el clúster:

- `yarn.node-labels.enabled:true`
- `yarn.node-labels.am.default-node-label-expression: 'ON_DEMAND'`

Si no especifica el límite `ON_DEMAND` y los parámetros máximos de nodos `CORE`, ambos parámetros se establecen de forma predeterminada en el límite máximo.

Si el máximo de nodos `CORE` es inferior al límite máximo, el escalado administrado usa el parámetro máximo de nodos `CORE` para dividir la asignación de capacidad entre los nodos `CORE` y `TASK`. Si establece el parámetro máximo de nodo `CORE` en un valor inferior o igual al parámetro de capacidad mínima, los nodos `CORE` permanecen estáticos con la capacidad máxima básica.

En los siguientes ejemplos, se muestra el escenario en el que se escalan las instancias bajo demanda en función de la demanda del proceso de aplicación y las instancias de Spot en función de la demanda del ejecutor.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Grupos de instancias Principal: 1 bajo demanda Tarea: 1 baja demanda	`UnitType`: instancias `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 10	Escala los nodos `ON_DEMAND` entre 1 y 20 nodos en función de la demanda de procesos de aplicación del clúster mediante el tipo de nodo `CORE` o `TASK`.

Estado inicial del clúster	Parámetros de escalado	Comportamiento del escalado
Flotas de instancias Principal: 1 bajo demanda Tarea: 1 baja demanda	`UnitType`: InstanceFleetUnits `MinimumCapacityUnits` : 1 `MaximumCapacityUnits` : 20 `MaximumOnDemandCapacityUnits` : 20 `MaximumCoreCapacityUnits` : 10	Escala los nodos `SPOT` en función de la demanda del ejecutor y de la capacidad restante disponible después de que Amazon EMR asigne los nodos `ON_DEMAND` . La suma de los nodos `SPOT` y `ON_DEMAND` solicitados no superará la `maximumCapacity` de 20. La suma de los nodos básicos bajo demanda y los nodos básicos de Spot solicitados no superará la `maximumCoreCapacity` de 10. Los nodos de Spot o bajo demanda adicionales utilizan el tipo de nodo `TASK`.

Descripción de las métricas de escalado administrado en Amazon EMR

Amazon EMR publica métricas de alta resolución con datos en una granularidad de un minuto cuando se habilita el escalado administrado para un clúster. Puede ver los eventos de cada inicio y finalización del cambio de tamaño controlados mediante el escalado gestionado con la consola Amazon EMR o la consola Amazon CloudWatch . CloudWatch las métricas son fundamentales para que funcione el escalado gestionado de Amazon EMR. Le recomendamos que supervise de cerca CloudWatch las métricas para asegurarse de que no falten datos. Para obtener más información sobre cómo configurar CloudWatch las alarmas para detectar las métricas faltantes, consulta Uso de CloudWatch las alarmas de Amazon. Para obtener más información sobre el uso de CloudWatch eventos con Amazon EMR, consulte Supervisar CloudWatch eventos.

Las siguientes métricas indican la capacidad actual o prevista de un clúster. Estas métricas solo están disponibles cuando el escalado administrado está habilitado. Para los clústeres compuestos por flotas de instancias, las métricas de capacidad del clúster se miden en `Units`. Para los clústeres compuestos por grupos de instancias, las métricas de capacidad del clúster se miden en `Nodes` o en `vCPU` en función del tipo de unidad utilizado en la política de escalado administrado.

Métrica	Descripción
• `TotalUnitsRequested` • `TotalNodesRequested` • `TotalVCPURequested`	El número total objetivo de units/nodes/vCPUs un clúster, determinado por el escalado administrado. Unidades: recuento
• `TotalUnitsRunning` • `TotalNodesRunning` • `TotalVCPURunning`	El número total actual de units/nodes/vCPUs disponibles en un clúster en ejecución. Cuando se solicita un cambio de tamaño del clúster, esta métrica se actualizará después de agregar o quitar las nuevas instancias del clúster. Unidades: recuento
• `CoreUnitsRequested` •	El número objetivo de CORE units/nodes/vCPUs en un clúster, determinado por el escalado gestionado.

Métrica	Descripción
• `CoreNodesRequested` • `CoreVCPURequested`	Unidades: recuento
• `CoreUnitsRunning` • `CoreNodesRunning` • `CoreVCPURunning`	El número actual de CORE units/nodes/vCPUs que se ejecutan en un clúster. Unidades: recuento
• `TaskUnitsRequested` • `TaskNodesRequested` • `TaskVCPURequested`	El número objetivo de TAREAS units/nodes/vCPUs en un clúster, determinado por el escalado gestionado. Unidades: recuento
• `TaskUnitsRunning` • `TaskNodesRunning` • `TaskVCPURunning`	El número actual de TAREAS units/nodes/vCPUs que se están ejecutando en un clúster. Unidades: recuento

Las siguientes métricas indican el estado de uso del clúster y las aplicaciones. Estas métricas están disponibles para todas las características de Amazon EMR, pero se publican con una resolución más alta con datos y una granularidad de un minuto cuando se habilita el escalado administrado para un clúster. Puede comparar las siguientes métricas con las métricas de capacidad del clúster de la tabla anterior para conocer las decisiones de escalado administrado.

Métrica	Descripción
`AppsCompleted`	

Métrica	Descripción
	El número de aplicaciones enviadas a YARN que se han completado. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
`AppsPending`	El número de aplicaciones enviadas a YARN que están en estado pendiente. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
`AppsRunning`	El número de aplicaciones enviadas a YARN que se están ejecutando. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
`ContainerAllocated`	El número de contenedores de recursos asignados porResourceManager. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
`ContainerPending`	El número de contenedores en la cola que aún no se han asignado. Caso de uso: monitorizar el progreso del clúster Unidades: recuento

Métrica	Descripción
`ContainerPendingRatio`	La relación entre los contenedores pendientes y los contenedores asignados (ContainerPendingRatio = ContainerPending / ContainerAllocated). Si Container Allocated = 0, entonces ContainerPendingRatio =Containe rPending. El valor de ContainerPendingRatio representa un número, no un porcentaje. Este valor es útil para escalar recursos del clúster en función del comportamiento de asignación de contenedores. Unidades: recuento
`HDFSUtilization`	El porcentaje de almacenamiento HDFS usado actualmente. Caso de uso: analizar el rendimiento del clúster Unidades: porcentaje
`IsIdle`	Indica que un clúster ya no está funcionando, pero sigue activo y acumulando cargos. Se establece en 1 si no se ejecuta ninguna tarea ni ningún trabajo; en caso contrario , se establece en 0. Este valor se comprueba a intervalo s de cinco minutos, y un valor de 1 indica que el clúster estaba inactivo cuando se comprobó, no que estuvo inactivo durante los cinco minutos. Para evitar falsos positivos, debe activar una alarma cuando este valor sea 1 durante más de una comprobación consecutiva de cinco minutos. Por ejemplo, puede activar una alarma cuando este valor sea 1 durante treinta minutos o más. Caso de uso: monitorizar el rendimiento del clúster Unidades: booleano

Métrica	Descripción
`MemoryAvailableMB`	La cantidad de memoria disponible para asignar. Caso de uso: monitorizar el progreso del clúster Unidades: recuento
`MRActiveNodes`	El número de nodos que actualmente ejecutan MapReduce tareas o trabajos. Equivalente a la métrica YARN `mapred.resourcemanager.NoOfActiveNodes` . Caso de uso: monitorizar el progreso del clúster Unidades: recuento
`YARNMemoryAvailablePercentage`	El porcentaje de memoria restante disponible para YARN (YARNMemoryAvailablePercentage = MemoryAvailable MB/ MemoryTotalMB). Este valor es útil para escalar recursos del clúster en función del uso de memoria de YARN. Unidades: porcentaje

Las siguientes métricas proporcionan información sobre los recursos utilizados por los contenedores y nodos de YARN. Estas métricas del administrador de recursos de YARN ofrecen información sobre los recursos que utilizan los contenedores y los nodos que se ejecutan en el clúster. Al comparar estas métricas con las métricas de capacidad del clúster de la tabla anterior, se obtiene una imagen más clara del impacto del escalado gestionado:

Métrica	Versiones asociadas	Descripción
`YarnContainersUsedMemoryGBSeconds`	Disponible para la etiqueta de lanzamiento 7.3.0 y versiones posteriores	La memoria consumida en el contenedor (* segundos) durante el período de publicación.

Métrica	Versiones asociadas	Descripción
		Unidades: GB * segundos
`YarnContainersTotalMemoryGBSeconds`	Disponible para la etiqueta de lanzamiento 7.3.0 y versiones posteriores	El contenido total de lana (* segundos) durante el período de publicación. Unidades: GB * segundos
`YarnContainersUsedVCPUSeconds`	Disponible para la etiqueta de lanzamiento 7.5.0 y superior	Los segundos de CPU virtual del contenedor consumidos durante el período de publicación. Unidades: CPU vCPU * segundos
`YarnContainersTotalVCPUSeconds`	Disponible para la etiqueta de lanzamiento 7.5.0 y versiones posteriores	El total de segundos de CPU virtual del contenedor durante el período de publicación. Unidades: CPU vCPU (* segundos)
`YarnNodesUsedMemoryGBSeconds`	Disponible para la etiqueta de lanzamiento 7.5.0 y versiones posteriores	La memoria del nodo consumida en segundos durante el período de publicación. Unidades: GB * segundos
`YarnNodesTotalMemoryGBSeconds`	Disponible para la etiqueta de lanzamiento 7.5.0 y superior	La memoria total del nodo (* segundos) durante el período de publicación. Unidades: GB * segundos

Métrica	Versiones asociadas	Descripción
`YarnNodesUsedVCPUSeconds`	Disponible para la etiqueta de lanzamiento 7.3.0 y versiones posteriores	Los segundos de CPU virtual del nodo consumidos durante el período de publicación. Unidades: CPU vCPU * segundos
`YarnNodesTotalVCPUSeconds`	Disponible para la etiqueta de lanzamiento 7.3.0 y versiones posteriores	El total de segundos de CPU virtual del nodo durante el período de publicación. Unidades: CPU virtual (* segundos)

Diagramación de métricas de escalado administrado

Puede diagramar las métricas para ver los patrones de carga de trabajo del clúster y las decisiones de escalado correspondientes tomadas por Escalado administrado de Amazon EMR, como se muestra en los pasos siguientes.

Para graficar las métricas de escalado gestionado en la consola CloudWatch

1. Abra la consola de CloudWatch.
2. En el panel de navegación, seleccione Amazon EMR. Puede buscar el identificador del clúster que desea monitorizar.
3. Desplácese hacia abajo hasta la métrica que desea representar gráficamente. Abra una métrica para mostrar el gráfico.
4. Para representar gráficamente una o varias métricas, seleccione la casilla de verificación junto a cada métrica.

En el siguiente ejemplo, se ilustra la actividad de Escalado administrado de Amazon EMR de un clúster. El gráfico muestra tres periodos de reducción de capacidad automática, que ahorran costos cuando hay una carga de trabajo menos activa.

Todas las métricas de capacidad y uso del clúster se publican a intervalos de un minuto. La información estadística adicional también se asocia a cada dato de un minuto, lo que le permite diagramar varias funciones como `Percentiles`, `Min`, `Max`, `Sum`, `Average`, `SampleCount`.

Por ejemplo, el siguiente gráfico muestra la misma métrica `YARNMemoryAvailablePercentage` en percentiles diferentes, P10, P50, P90, P99, junto con `Sum`, `Average`, `Min`, `SampleCount`.

Uso del escalado automático con una política personalizada para grupos de instancias en Amazon EMR

El escalado automático con una política personalizada en las versiones 4.0 y superiores de Amazon EMR le permite escalar programáticamente los nodos principales y los nodos de tareas en función de una CloudWatch métrica y otros parámetros que especifique en una política de escalado. El escalado automático con una política personalizada está disponible con la configuración de grupos de instancias, pero no con las flotas de instancias. Para obtener más información sobre los grupos de instancias y las flotas de instancias, consulte Creación de un clúster de Amazon EMR con flotas de instancias o grupos de instancias uniformes.

> **Note**
>
> Para utilizar la característica de escalado automático con una política personalizada en Amazon EMR, debe establecer en `true` el parámetro `VisibleToAllUsers` al crear un clúster. Para obtener más información, consulte SetVisibleToAllUsers.

La política de escalado forma parte de la configuración de un grupo de instancias. Puede especificar una política durante la configuración de un grupo de instancias o modificando un grupo de instancias en un clúster existente, incluso cuando dicho grupo de instancias está activo. Cada grupo de instancia de un clúster, excepto el grupo de instancias principales, puede tener su propia política de escalado, que se compone de reglas de escalado horizontal y reducción horizontal. Las reglas de escalado ascendente y descendente se pueden configurar de manera independiente, con distintos parámetros para cada regla.

Puede configurar políticas de escalado con la AWS Management Console AWS CLI, la o la API de Amazon EMR. Cuando utiliza la API AWS CLI o Amazon EMR, especifica la política de escalado en formato JSON. Además, si utiliza la API Amazon EMR AWS CLI o la API, puede especificar métricas personalizadas CloudWatch . Las métricas personalizadas no se pueden seleccionar con la AWS

Management Console. Al crear inicialmente una política de escalado con la consola, se preconfigura una política predeterminada adecuada para muchas aplicaciones para ayudarle a comenzar. Puede eliminar o modificar las reglas predeterminadas.

Si bien el escalado automático te permite ajustar la capacidad del clúster de EMR on-the-fly, debes tener en cuenta los requisitos de carga de trabajo básicos y planificar las configuraciones de los nodos y grupos de instancias. Para obtener más información, consulte Directrices de configuración del clúster.

> **Note**
>
> Para la mayoría de las cargas de trabajo, es deseable la configuración de las reglas de escalado ascendente y descendente para optimizar el uso de los recursos. Definir una regla sin la otra significa que tendrá que cambiar manualmente el tamaño del recuento de instancias después de una actividad de escalado. En otras palabras, esto configura una política de escalado ascendente o descendente automática "unidireccional" con un restablecimiento manual.

Creación del rol de IAM; para el escalado automático

El escalado automático en Amazon EMR requiere un rol de IAM con permisos para agregar y terminar instancias cuando se activan las actividades de escalado. Para este fin se dispone de un rol predeterminado, `EMR_AutoScaling_DefaultRole`, configurado con la política de confianza y la política de roles adecuadas. Cuando crea un clúster con una política de escalado por primera vez con AWS Management Console, Amazon EMR crea el rol predeterminado y adjunta la política administrada predeterminada para los permisos,. `AmazonElasticMapReduceforAutoScalingRole`

Al crear un clúster con una política de escalado automático con el AWS CLI, primero debe asegurarse de que existe el rol de IAM predeterminado o de que tiene un rol de IAM personalizado con una política adjunta que proporcione los permisos adecuados. Para crear el rol predeterminado, puede ejecutar el comando `create-default-roles` antes de crear un clúster. A continuación, puede especificar la opción `--auto-scaling-role EMR_AutoScaling_DefaultRole` al crear el clúster. También puede crear un rol de escalado automático personalizado y luego especificarlo al crear un clúster, por ejemplo, `--auto-scaling-role` *`MyEMRAutoScalingRole`*. Si crea un rol de escalado automático personalizado para Amazon EMR, le recomendamos que base las políticas de permisos de dicho rol en la política administrada. Para obtener más información, consulte

Configuración de los roles de servicio de IAM de los permisos de Amazon EMR para los servicios y recursos de AWS.

Descripción de las reglas de escalado automático

Cuando una regla de escalado horizontal desencadena una actividad de escalado para un grupo de instancias, EC2 las instancias de Amazon se añaden al grupo de instancias de acuerdo con tus reglas. Aplicaciones como Apache Spark, Apache Hive y Presto pueden usar los nodos nuevos tan pronto como la EC2 instancia de Amazon entre en el `InService` estado. También puede configurar una regla de escalado descendente que termina instancias y elimina nodos. Para obtener más información sobre el ciclo de vida de las EC2 instancias de Amazon que se escalan automáticamente, consulte el ciclo de vida de Auto Scaling en la Guía del usuario de Amazon EC2 Auto Scaling.

Puede configurar la forma en que un clúster termina las EC2 instancias de Amazon. Puedes elegir entre finalizar la facturación en el límite de EC2 horas de instancia de Amazon o al finalizar la tarea. Esta configuración se aplica tanto al escalado automático como a las operaciones manuales de cambio de tamaño. Para obtener más información acerca de esta configuración, consulte Opciones de reducción vertical de clústeres para clústeres de Amazon EMR.

Los siguientes parámetros para cada regla de una política determinan el comportamiento de escalado automático.

> **Note**
>
> Los parámetros que se muestran aquí se basan en los AWS Management Console de Amazon EMR. Al utilizar la API AWS CLI o Amazon EMR, hay disponibles opciones de configuración avanzada adicionales. Para obtener más información sobre las opciones avanzadas, consulte SimpleScalingPolicyConfigurationla referencia de la API de Amazon EMR.

- Número máximo de instancias y número mínimo de instancias. La restricción Máximo de instancias especifica el número máximo de EC2 instancias de Amazon que pueden estar en el grupo de instancias y se aplica a todas las reglas de escalado horizontal. Del mismo modo, la restricción de instancias mínimas especifica el número mínimo de EC2 instancias de Amazon y se aplica a todas las reglas de escalado interno.
- El Rule name (Nombre de la regla), que debe ser único en la política.

- El ajuste de escalado, que determina el número de EC2 instancias que se deben añadir (en el caso de las reglas de escalamiento horizontal) o finalizar (en el caso de las reglas de escalado interno) durante la actividad de escalado que desencadena la regla.
- La CloudWatch métrica, que se vigila en busca de una condición de alarma.
- Un operador de comparación, que se utiliza para comparar la CloudWatch métrica con el valor umbral y determinar una condición de activación.
- Un período de evaluación, en incrementos de cinco minutos, durante el cual la CloudWatch métrica debe estar en una condición de activación antes de que se active la actividad de escalado.
- Un Cooldown period (período de recuperación) en segundos, que determina la cantidad de tiempo que debe transcurrir entre una actividad de escalado iniciada por una regla y el inicio de la próxima actividad de escalado, con independencia de la regla que la activa. Cuando un grupo de instancias finaliza una actividad de escalado y alcanza su estado posterior a la escalabilidad, el período de enfriamiento brinda una oportunidad para que las CloudWatch métricas que podrían desencadenar actividades de escalado posteriores se estabilicen. Para obtener más información, consulte los tiempos de reutilización de Auto Scaling en la Guía del usuario de Amazon EC2 Auto Scaling.

Consideraciones y limitaciones

- CloudWatch Las métricas de Amazon son fundamentales para que funcione el escalado automático de Amazon EMR. Te recomendamos que supervises de cerca CloudWatch las estadísticas de Amazon para asegurarte de que no falten datos. Para obtener más información sobre cómo configurar CloudWatch las alarmas de Amazon para detectar las métricas faltantes, consulta Uso de CloudWatch las alarmas de Amazon.
- La sobreutilización de los volúmenes de EBS puede provocar problemas de Escalado administrado. Le recomendamos que supervise de cerca el uso de los volúmenes de EBS para asegurarse de que el volumen de EBS esté por debajo del 90 % de utilización. Consulte Almacenamiento de instancias para obtener información sobre cómo especificar volúmenes de EBS adicionales.
- El escalado automático con una política personalizada en las versiones 5.18 a 5.28 de Amazon EMR puede experimentar un error de escalado debido a la ausencia intermitente de datos en las métricas de Amazon. CloudWatch Para mejorar el escalado automático, le recomendamos que utilice las versiones más recientes de Amazon EMR. También puede ponerse en contacto con AWS Support para solicitar un parche si tiene que utilizar una versión de Amazon EMR que esté entre la 5.18 y la 5.28.

Utilizándolo para configurar el escalado automático AWS Management Console

Al crear un clúster, se configura una política de escalado para grupos de instancias con las opciones avanzadas de configuración de clúster. También puede crear o modificar una política de escalado para un grupo de instancias en servicio modificando los grupos de instancias en la configuración de Hardware de un clúster existente.

1. Vaya hasta la nueva consola de Amazon EMR y seleccione Ir a la consola antigua en el panel de navegación lateral. Para más información sobre lo que puede esperar al cambiar a la consola antigua, consulte Uso de la consola antigua.
2. Si está creando un clúster, en la consola de Amazon EMR, seleccione Crear clúster, seleccione Ir a las opciones avanzadas, seleccione las opciones en Paso 1: Software y pasos y, a continuación, acceda a Paso 2: Configuración de hardware.

   - o bien -

   Si está modificando un grupo de instancias en un clúster en ejecución, seleccione el clúster en la lista de clústeres y, a continuación, expanda la sección Hardware.
3. En la sección Opción de escalado y aprovisionamiento del clúster, seleccione Habilitar el escalado del clúster. A continuación, seleccione Crear una política de escalado automático personalizada.

   En la tabla Políticas de escalado automático personalizadas, haga clic en el icono de lápiz que aparece en la fila del grupo de instancias que desea configurar. Se abre la pantalla de reglas de Auto Scaling.
4. Escriba el valor de Maximum instances (Número máximo de instancias) que desea que contenga el grupo de instancias después del escalado ascendente y el valor de Minimum instances (Número mínimo de instancias) que desea que contenga el grupo de instancias después del escalado descendente.
5. Haga clic en el icono de lápiz para editar los parámetros de regla, haga clic en la X para eliminar una regla de la política y, a continuación, haga clic en Add rule (Añadir regla) para añadir reglas adicionales.
6. Elija los parámetros de la configuración, tal como se describió anteriormente en este tema. Para obtener descripciones de las CloudWatch métricas disponibles para Amazon EMR, consulte las métricas y dimensiones de Amazon EMR en la Guía del usuario de Amazon. CloudWatch

Utilizándolo AWS CLI para configurar el escalado automático

Puede usar AWS CLI los comandos de Amazon EMR para configurar el escalado automático al crear un clúster y al crear un grupo de instancias. Puede utilizar una sintaxis abreviada, especificando la configuración JSON insertada dentro de los comandos pertinentes o puede hacer referencia a un archivo que contenga la configuración JSON. También puede aplicar una política de escalado automático a un grupo de instancias existente y eliminar la política de escalado automático que se aplicó anteriormente. Además, puede recuperar detalles de una configuración de política de escalado desde un clúster en ejecución.

> **Important**
>
> Al crear un clúster con una política de escalado automático, es necesario utilizar el comando `--auto-scaling-role MyAutoScalingRole` para especificar el rol de IAM para el escalado automático. El rol predeterminado es `EMR_AutoScaling_DefaultRole` y se puede crear mediante el comando `create-default-roles`. El rol solo se puede añadir cuando se crea el clúster y no se puede añadir a un clúster existente.

Para obtener una descripción detallada de los parámetros disponibles al configurar una política de escalado automático, consulte la referencia PutAutoScalingPolicyde la API de Amazon EMR.

Creación de un clúster con una política de escalado automático aplicada a un grupo de instancias

Puede especificar una configuración de escalado automático dentro de la opción `--instance-groups` del comando `aws emr create-cluster`. El siguiente ejemplo ilustra un comando create-cluster donde se proporciona una política insertada de escalado automático para el grupo de instancias secundarias. El comando crea una configuración de escalado equivalente a la política de escalado horizontal predeterminada que aparece al crear una política de escalado automático con AWS Management Console Amazon EMR. Por razones de brevedad, no se muestra una política de escalado descendente. No recomendamos la creación de una regla de escalado ascendente sin una regla de escalado descendente.

```
aws emr create-cluster --release-label emr-5.2.0 --service-role
 EMR_DefaultRole --ec2-attributes InstanceProfile=EMR_EC2_DefaultRole
 --auto-scaling-role EMR_AutoScaling_DefaultRole  --instance-groups
 Name=MyMasterIG,InstanceGroupType=MASTER,InstanceType=m5.xlarge,InstanceCount=1
 'Name=MyCoreIG,InstanceGroupType=CORE,InstanceType=m5.xlarge,InstanceCount=2,AutoScalingPolicy
scale-out,Description=Replicates the default scale-out rule in the
 console.,Action={SimpleScalingPolicyConfiguration={AdjustmentType=CHANGE_IN_CAPACITY,ScalingAd
```

```
ElasticMapReduce,Period=300,Statistic=AVERAGE,Threshold=15,Unit=PERCENT,Dimensions=[{Key=JobFlo
```

El siguiente comando ilustra el uso de la línea de comandos para proporcionar la definición de política de escalado automático como parte de un archivo de configuración de grupo de instancias denominado *instancegroupconfig.json*.

```
aws emr create-cluster --release-label emr-5.2.0 --service-role EMR_DefaultRole --ec2-
attributes InstanceProfile=EMR_EC2_DefaultRole --instance-groups file://your/path/to/
instancegroupconfig.json --auto-scaling-role EMR_AutoScaling_DefaultRole
```

Con el contenido del archivo de configuración siguiente:

```
[
{
  "InstanceCount": 1,
  "Name": "MyMasterIG",
  "InstanceGroupType": "MASTER",
  "InstanceType": "m5.xlarge"
},
{
  "InstanceCount": 2,
  "Name": "MyCoreIG",
  "InstanceGroupType": "CORE",
  "InstanceType": "m5.xlarge",
  "AutoScalingPolicy":
    {
     "Constraints":
      {
       "MinCapacity": 2,
       "MaxCapacity": 10
      },
     "Rules":
     [
      {
       "Name": "Default-scale-out",
       "Description": "Replicates the default scale-out rule in the console for YARN
 memory.",
       "Action":{
        "SimpleScalingPolicyConfiguration":{
          "AdjustmentType": "CHANGE_IN_CAPACITY",
          "ScalingAdjustment": 1,
```

```
          "CoolDown": 300
        }
       },
       "Trigger":{
        "CloudWatchAlarmDefinition":{
          "ComparisonOperator": "LESS_THAN",
          "EvaluationPeriods": 1,
          "MetricName": "YARNMemoryAvailablePercentage",
          "Namespace": "AWS/ElasticMapReduce",
          "Period": 300,
          "Threshold": 15,
          "Statistic": "AVERAGE",
          "Unit": "PERCENT",
          "Dimensions":[
             {
               "Key" : "JobFlowId",
               "Value" : "${emr.clusterId}"
             }
          ]
        }
       }
      }
     ]
   }
}
]
```

Agregar un grupo de instancias con una política de escalado automático a un clúster

Puede especificar una configuración de política de escalado con la opción `--instance-groups` con el comando `add-instance-groups` de la misma forma que al utilizar `create-cluster`. En el siguiente ejemplo se utiliza una referencia a un archivo JSON, *`instancegroupconfig.json`*, con la configuración de grupo de instancias.

```
aws emr add-instance-groups --cluster-id j-1EKZ3TYEVF1S2 --instance-groups file://your/
path/to/instancegroupconfig.json
```

Aplicar una política de escalado automático a un grupo de instancias existente o modificar una política aplicada

Utilice el comando `aws emr put-auto-scaling-policy` para aplicar una política de escalado automático a un grupo de instancias existente. El grupo de instancias debe formar parte de un clúster

que utilice el rol de IAM; de escalado automático. En el siguiente ejemplo se utiliza una referencia a un archivo JSON, *autoscaleconfig.json*, que especifica la configuración de política de escalado.

```
aws emr put-auto-scaling-policy --cluster-id j-1EKZ3TYEVF1S2 --instance-group-
id ig-3PLUZBA6WLS07 --auto-scaling-policy file://your/path/to/autoscaleconfig.json
```

El contenido del archivo `autoscaleconfig.json`, que define la misma regla de escalado ascendente, tal y como se muestra en el ejemplo anterior, se muestra a continuación.

```
{
          "Constraints": {
                  "MaxCapacity": 10,
                  "MinCapacity": 2
          },
          "Rules": [{
                  "Action": {
                          "SimpleScalingPolicyConfiguration": {
                                  "AdjustmentType": "CHANGE_IN_CAPACITY",
                                  "CoolDown": 300,
                                  "ScalingAdjustment": 1
                          }
                  },
                  "Description": "Replicates the default scale-out rule in the console
 for YARN memory",
                  "Name": "Default-scale-out",
                  "Trigger": {
                          "CloudWatchAlarmDefinition": {
                                  "ComparisonOperator": "LESS_THAN",
                                  "Dimensions": [{
                                          "Key": "JobFlowId",
                                          "Value": "${emr.clusterID}"
                                  }],
                                  "EvaluationPeriods": 1,
                                  "MetricName": "YARNMemoryAvailablePercentage",
                                  "Namespace": "AWS/ElasticMapReduce",
                                  "Period": 300,
                                  "Statistic": "AVERAGE",
                                  "Threshold": 15,
                                  "Unit": "PERCENT"
                          }
                  }
```

```
        }]
  }
```

Eliminación de una política de escalado automático desde un grupo de instancias

```
aws emr remove-auto-scaling-policy --cluster-id j-1EKZ3TYEVF1S2 --instance-group-
id ig-3PLUZBA6WLS07
```

Recuperación de una configuración de política de escalado automático

El `describe-cluster` comando recupera la configuración de políticas del bloque. InstanceGroup Por ejemplo, el comando siguiente recupera la configuración para el clúster con un ID de clúster de `j-1CWOHP4PI30VJ`.

```
aws emr describe-cluster --cluster-id j-1CWOHP4PI30VJ
```

El comando produce el siguiente resultado de ejemplo.

```
{
    "Cluster": {
        "Configurations": [],
        "Id": "j-1CWOHP4PI30VJ",
        "NormalizedInstanceHours": 48,
        "Name": "Auto Scaling Cluster",
        "ReleaseLabel": "emr-5.2.0",
        "ServiceRole": "EMR_DefaultRole",
        "AutoTerminate": false,
        "TerminationProtected": true,
        "MasterPublicDnsName": "ec2-54-167-31-38.compute-1.amazonaws.com",
        "LogUri": "s3n://aws-logs-232939870606-us-east-1/elasticmapreduce/",
        "Ec2InstanceAttributes": {
            "Ec2KeyName": "performance",
            "AdditionalMasterSecurityGroups": [],
            "AdditionalSlaveSecurityGroups": [],
            "EmrManagedSlaveSecurityGroup": "sg-09fc9362",
            "Ec2AvailabilityZone": "us-east-1d",
            "EmrManagedMasterSecurityGroup": "sg-0bfc9360",
            "IamInstanceProfile": "EMR_EC2_DefaultRole"
        },
```

```
        "Applications": [
            {
                "Name": "Hadoop",
                "Version": "2.7.3"
            }
        ],
        "InstanceGroups": [
            {
                "AutoScalingPolicy": {
                    "Status": {
                        "State": "ATTACHED",
                        "StateChangeReason": {
                            "Message": ""
                        }
                    },
                    "Constraints": {
                        "MaxCapacity": 10,
                        "MinCapacity": 2
                    },
                    "Rules": [
                        {
                            "Name": "Default-scale-out",
                            "Trigger": {
                                "CloudWatchAlarmDefinition": {
                                    "MetricName": "YARNMemoryAvailablePercentage",
                                    "Unit": "PERCENT",
                                    "Namespace": "AWS/ElasticMapReduce",
                                    "Threshold": 15,
                                    "Dimensions": [
                                        {
                                            "Key": "JobFlowId",
                                            "Value": "j-1CWOHP4PI30VJ"
                                        }
                                    ],
                                    "EvaluationPeriods": 1,
                                    "Period": 300,
                                    "ComparisonOperator": "LESS_THAN",
                                    "Statistic": "AVERAGE"
                                }
                            },
                            "Description": "",
                            "Action": {
                                "SimpleScalingPolicyConfiguration": {
                                    "CoolDown": 300,
```

```
                                "AdjustmentType": "CHANGE_IN_CAPACITY",
                                "ScalingAdjustment": 1
                            }
                        }
                    },
                    {
                        "Name": "Default-scale-in",
                        "Trigger": {
                            "CloudWatchAlarmDefinition": {
                                "MetricName": "YARNMemoryAvailablePercentage",
                                "Unit": "PERCENT",
                                "Namespace": "AWS/ElasticMapReduce",
                                "Threshold": 75,
                                "Dimensions": [
                                    {
                                        "Key": "JobFlowId",
                                        "Value": "j-1CWOHP4PI30VJ"
                                    }
                                ],
                                "EvaluationPeriods": 1,
                                "Period": 300,
                                "ComparisonOperator": "GREATER_THAN",
                                "Statistic": "AVERAGE"
                            }
                        },
                        "Description": "",
                        "Action": {
                            "SimpleScalingPolicyConfiguration": {
                                "CoolDown": 300,
                                "AdjustmentType": "CHANGE_IN_CAPACITY",
                                "ScalingAdjustment": -1
                            }
                        }
                    }
                ]
            },
            "Configurations": [],
            "InstanceType": "m5.xlarge",
            "Market": "ON_DEMAND",
            "Name": "Core - 2",
            "ShrinkPolicy": {},
            "Status": {
                "Timeline": {
                    "CreationDateTime": 1479413437.342,
```

```
                        "ReadyDateTime": 1479413864.615
                    },
                    "State": "RUNNING",
                    "StateChangeReason": {
                        "Message": ""
                    }
                },
                "RunningInstanceCount": 2,
                "Id": "ig-3M16XBE8C3PH1",
                "InstanceGroupType": "CORE",
                "RequestedInstanceCount": 2,
                "EbsBlockDevices": []
            },
            {
                "Configurations": [],
                "Id": "ig-OP62I28NSE8M",
                "InstanceGroupType": "MASTER",
                "InstanceType": "m5.xlarge",
                "Market": "ON_DEMAND",
                "Name": "Master - 1",
                "ShrinkPolicy": {},
                "EbsBlockDevices": [],
                "RequestedInstanceCount": 1,
                "Status": {
                    "Timeline": {
                        "CreationDateTime": 1479413437.342,
                        "ReadyDateTime": 1479413752.088
                    },
                    "State": "RUNNING",
                    "StateChangeReason": {
                        "Message": ""
                    }
                },
                "RunningInstanceCount": 1
            }
        ],
        "AutoScalingRole": "EMR_AutoScaling_DefaultRole",
        "Tags": [],
        "BootstrapActions": [],
        "Status": {
            "Timeline": {
                "CreationDateTime": 1479413437.339,
                "ReadyDateTime": 1479413863.666
            },
```

```
            "State": "WAITING",
            "StateChangeReason": {
                "Message": "Cluster ready after last step completed."
            }
        }
    }
}
```

Cambio manual del tamaño de un clúster de Amazon EMR en ejecución

Puede añadir y eliminar instancias de los grupos de instancias principales y de tareas y de las flotas de instancias de un clúster en ejecución con la AWS CLI API AWS Management Console, o Amazon EMR. Si un clúster utiliza grupos de instancias, es necesario cambiar de forma explícita el recuento de instancias. Si el clúster utiliza flotas de instancias, puede cambiar las unidades de destino en las instancias bajo demanda y las instancias de spot. A continuación, la flota de instancias añadirá y eliminará instancias para satisfacer los requisitos del nuevo destino. Para obtener más información, consulte Opciones de flota de instancias. Las aplicaciones pueden usar EC2 instancias de Amazon recién aprovisionadas para alojar nodos en cuanto las instancias estén disponibles. Cuando se eliminan instancias, Amazon EMR finaliza las tareas de forma que no se interrumpan los trabajos y se protejan frente a la pérdida de datos. Para obtener más información, consulte Terminación al completar las tareas.

Cambiar el tamaño de un clúster con la consola

Puede utilizar la consola de Amazon EMR para cambiar el tamaño de un clúster en ejecución.

Console

Para cambiar el recuento de instancias de un clúster existente con la nueva consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y seleccione el clúster que desee actualizar. El clúster debe estar en ejecución; no puede cambiar el tamaño de un clúster que se esté aprovisionando o se haya terminado.
3. En la pestaña Instancias de la página de detalles del clúster, observe el panel Grupos de instancias.

4. Para cambiar el tamaño de un grupo de instancias existente, selecciona el botón de opción situado junto al grupo de instancias principales o de tarea cuyo tamaño desee cambiar y, a continuación, seleccione Cambiar el tamaño del grupo de instancias. Especifique el nuevo número de instancias para el grupo de instancias y, a continuación, seleccione Cambiar tamaño.

> **Note**
>
> Si decide reducir el tamaño de un grupo de instancias en ejecución, Amazon EMR seleccionará de forma inteligente las instancias que desee eliminar del grupo para reducir al mínimo la pérdida de datos. Para tener un control más detallado de la acción de cambio de tamaño, puede seleccionar el ID del grupo de instancias, seleccionar las instancias que desea eliminar y, a continuación, utilizar la opción Terminar. Para obtener más información sobre el comportamiento inteligente de reducción vertical, consulte Opciones de reducción vertical de clústeres para clústeres de Amazon EMR.

5. Si desea cancelar la acción de cambio de tamaño, puede seleccionar el botón de opción de un grupo de instancias con el estado Cambiando el tamaño y, a continuación, seleccionar Detener el cambio de tamaño en la lista de acciones.
6. Para agregar uno o más grupos de instancias de tarea al clúster en respuesta al aumento de la carga de trabajo, seleccione Agregar grupo de instancias de tareas en la lista de acciones. Elige el tipo de EC2 instancia de Amazon, introduce el número de instancias del grupo de tareas y, a continuación, selecciona Añadir grupo de instancias de tareas para volver al panel Grupos de instancias de tu clúster.

Cuando se realiza un cambio en el número de nodos, se actualiza el Status (Estado) de las actualizaciones del grupo de instancias. Cuando finaliza el cambio solicitado, el Status (Estado) es Running (En ejecución).

Cambie el tamaño de un clúster con la AWS CLI

Puede utilizar el AWS CLI para cambiar el tamaño de un clúster en ejecución. Puede aumentar o reducir el número de nodos de tareas y puede aumentar el número de nodos secundarios en un clúster en ejecución. También es posible cerrar una instancia del grupo de instancias principal con la API AWS CLI o la misma. Esto debe hacerse con precaución. Apagar una instancia en el grupo

de instancias secundarias conlleva un riesgo de pérdida de datos y la instancia no se sustituye automáticamente.

Además de cambiar el tamaño de los grupos se instancias secundarias y de tareas, también puede agregar uno o varios grupos de instancias de tareas a un clúster en ejecución con la AWS CLI.

Para cambiar el tamaño de un clúster cambiando el recuento de instancias con la AWS CLI

Puede añadir instancias al grupo principal o al grupo de tareas y eliminar instancias del grupo de tareas con el AWS CLI `modify-instance-groups` subcomando del `InstanceCount` parámetro. Para agregar instancias a los grupos de instancias principales o de tareas, aumente el `InstanceCount`. Para reducir el número de instancias en el grupo de tareas, reduzca el `InstanceCount`. Al cambiar el recuento de instancias del grupo de tareas a 0 se eliminan todas las instancias, pero no el grupo de instancias.

- Para aumentar el número de instancias en el grupo de instancias de tareas de 3 a 4, escribe el siguiente comando y *`ig-31JXXXXXXBT0`* sustitúyelo por el ID del grupo de instancias.

  ```
  aws emr modify-instance-groups --instance-groups
   InstanceGroupId=ig-31JXXXXXXBT0,InstanceCount=4
  ```

  Para recuperar el `InstanceGroupId`, utilice el subcomando `describe-cluster`. La salida es un objeto JSON denominado `Cluster` que contiene el ID de cada grupo de instancias. Para utilizar este comando, necesita el ID del clúster (que puede recuperar con el comando `aws emr list-clusters` o la consola). Para recuperar el ID del grupo de instancias, escribe el siguiente comando y *`j-2AXXXXXXGAPLF`* sustitúyelo por el ID del clúster.

  ```
  aws emr describe-cluster --cluster-id j-2AXXXXXXGAPLF
  ```

  Con el AWS CLI, también puedes terminar una instancia del grupo de instancias principal con el `--modify-instance-groups` subcomando.

  > **Warning**
  >
  > Debe prestarse especial atención cuando se especifique `EC2InstanceIdsToTerminate`. Las instancias se terminan de forma inmediata, independientemente del estado de las aplicaciones que se ejecutan en ellas y de que la instancia no se sustituya automáticamente. Esto ocurre independientemente de la configuración de Scale down behavior (Comportamiento de escalado descendente) del

clúster. Terminar una instancia de esta forma conlleva el riesgo de pérdida de datos y de un comportamiento del clúster imprevisible.

Para terminar una instancia específica, necesitas el ID del grupo de instancias (devuelto por el `aws emr describe-cluster --cluster-id` subcomando) y el ID de la instancia (devuelto por el `aws emr list-instances --cluster-id` subcomando). Escribe el siguiente comando, sustitúyelo por el ID del grupo de instancias y *ig-6RXXXXXX07SA* sustitúyelo por el ID de la instancia. *i-f9XXXXf2*

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-6RXXXXXX07SA,EC2InstanceIdsToTerminate=i-f9XXXXf2
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Para cambiar el tamaño de un clúster añadiendo grupos de instancias de tareas con la AWS CLI

Con el AWS CLI, puedes añadir de 1 a 48 grupos de instancias de tareas a un clúster con el `--add-instance-groups` subcomando. Los grupos de instancias de tareas solo se pueden agregar a un clúster que contenga un grupo de instancias principales y un grupo de instancias de núcleo. Al utilizar el subcomando AWS CLI, puede añadir hasta cinco grupos de instancias de tareas cada vez que utilice el `--add-instance-groups` subcomando.

1. Para añadir un único grupo de instancias de tareas a un clúster, escribe el siguiente comando y *j-JXBXXXXXX37R* sustitúyelo por el ID del clúster.

```
aws emr add-instance-groups --cluster-id j-JXBXXXXXX37R --instance-groups
 InstanceCount=6,InstanceGroupType=task,InstanceType=m5.xlarge
```

2. Para añadir varios grupos de instancias de tareas a un clúster, escribe el siguiente comando y *j-JXBXXXXXX37R* sustitúyelo por el ID del clúster. Puede añadir hasta cinco grupos de instancias de tareas en un solo comando.

```
aws emr add-instance-groups --cluster-id j-JXBXXXXXX37R --instance-
groups InstanceCount=6,InstanceGroupType=task,InstanceType=m5.xlarge
 InstanceCount=10,InstanceGroupType=task,InstanceType=m5.xlarge
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Interrupción de un cambio de tamaño

Si utiliza la versión 4.1.0 o posterior de Amazon EMR, puede realizar un cambio de tamaño en medio de una operación de cambio de tamaño existente. Además, puede parar una solicitud de cambio de tamaño enviada previamente o enviar una nueva solicitud para anular una solicitud anterior sin necesidad de esperar a que finalice. También puede detener un cambio de tamaño existente desde la consola o a través de la llamada a la API `ModifyInstanceGroups` con el recuento actual como el recuento de destino del clúster.

La siguiente captura de pantalla muestra un grupo de instancias de tareas que se está cambiando de tamaño, pero que puede interrumpirse eligiendo Stop (Detener).

Para interrumpir un cambio de tamaño con el AWS CLI

Puede utilizar el AWS CLI para detener un cambio de tamaño con el `modify-instance-groups` subcomando. Supongamos que tiene seis instancias en un grupo de instancias y que desea aumentarlo a 10. Más tarde decide que desea cancelar esta solicitud:

- La solicitud inicial:

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-myInstanceGroupId,InstanceCount=10
```

  La segunda solicitud para detener la primera solicitud:

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-myInstanceGroupId,InstanceCount=6
```

> **Note**
>
> Dado que este proceso es asíncrono, es posible que vea que los recuentos de instancias cambian con respecto las solicitudes de la API previas antes de que se realicen las solicitudes siguientes. En el caso de disminución, es posible que si tiene el trabajo en

ejecución en los nodos, el grupo de instancias podría no reducirse hasta que los nodos hayan completado su trabajo.

Estado de suspensión

Un grupo de instancias pasa a estado suspendido si se encuentra demasiados errores al intentar iniciar los nuevos nodos del clúster. Por ejemplo, si los nodos nuevos devuelven error repetidamente al llevar a cabo acciones de arranque, el grupo de instancias pasa al estado SUSPENDED, en lugar de aprovisionar de forma continua nuevos nodos. Después de haber resuelto el error subyacente, restablezca el número deseado de nodos en el grupo de instancias del clúster y, a continuación, el grupo de instancias reanuda la asignación de nodos. La modificación de una instancia grupo indica a Amazon EMR que vuelva a intentar aprovisionar los nodos. Ningún nodo en ejecución se reinicia ni se termina.

En el AWS CLI, el `list-instances` subcomando devuelve todas las instancias y sus estados, al igual que el subcomando. `describe-cluster` Si Amazon EMR detecta un error con un grupo de instancias, cambia el estado del grupo a `SUSPENDED`.

Para restablecer un clúster en estado SUSPENDIDO con el AWS CLI

Escriba el subcomando `describe-cluster` con el parámetro `--cluster-id` para ver el estado de las instancias en el clúster.

- Para ver la información sobre todas las instancias y grupos de instancias de un clúster, escribe el siguiente comando y *`j-3KVXXXXXXY7UG`* sustitúyelo por el ID del clúster.

  ```
  aws emr describe-cluster --cluster-id j-3KVXXXXXXY7UG
  ```

  La salida muestra información sobre los grupos de la instancia y el estado de las instancias:

  ```
  {
      "Cluster": {
          "Status": {
              "Timeline": {
                  "ReadyDateTime": 1413187781.245,
                  "CreationDateTime": 1413187405.356
              },
              "State": "WAITING",
              "StateChangeReason": {
  ```

```
                "Message": "Waiting after step completed"
            }
        },
        "Ec2InstanceAttributes": {
            "Ec2AvailabilityZone": "us-west-2b"
        },
        "Name": "Development Cluster",
        "Tags": [],
        "TerminationProtected": false,
        "RunningAmiVersion": "3.2.1",
        "NormalizedInstanceHours": 16,
        "InstanceGroups": [
            {
                "RequestedInstanceCount": 1,
                "Status": {
                    "Timeline": {
                        "ReadyDateTime": 1413187775.749,
                        "CreationDateTime": 1413187405.357
                    },
                    "State": "RUNNING",
                    "StateChangeReason": {
                        "Message": ""
                    }
                },
                "Name": "MASTER",
                "InstanceGroupType": "MASTER",
                "InstanceType": "m5.xlarge",
                "Id": "ig-3ETXXXXXXFYV8",
                "Market": "ON_DEMAND",
                "RunningInstanceCount": 1
            },
            {
                "RequestedInstanceCount": 1,
                "Status": {
                    "Timeline": {
                        "ReadyDateTime": 1413187781.301,
                        "CreationDateTime": 1413187405.357
                    },
                    "State": "RUNNING",
                    "StateChangeReason": {
                        "Message": ""
                    }
                },
                "Name": "CORE",
```

```
                "InstanceGroupType": "CORE",
                "InstanceType": "m5.xlarge",
                "Id": "ig-3SUXXXXXXQ9ZM",
                "Market": "ON_DEMAND",
                "RunningInstanceCount": 1
            }
...
}
```

Para ver información sobre un grupo de instancias determinado, escriba el subcomando `list-instances` con los parámetros `--cluster-id` e `--instance-group-types`. Puede consultar información para los grupos principales, de núcleo o de tarea.

```
aws emr list-instances --cluster-id j-3KVXXXXXXY7UG --instance-group-types "CORE"
```

Utilice el subcomando `modify-instance-groups` con el parámetro `--instance-groups` para restablecer un clúster en el estado `SUSPENDED`. El subcomando `describe-cluster` devuelve el id del grupo de instancias.

```
aws emr modify-instance-groups --instance-groups
 InstanceGroupId=ig-3SUXXXXXXQ9ZM,InstanceCount=3
```

Consideraciones a la hora de reducir el tamaño del clúster

Si decide reducir el tamaño de un clúster en ejecución, tenga en cuenta el comportamiento y las prácticas recomendadas de Amazon EMR siguientes:

- Para reducir el impacto en los trabajos en curso, Amazon EMR selecciona de forma inteligente las instancias que se van a eliminar. Para más información sobre el comportamiento de reducción vertical del clúster, consulte Terminación al completar las tareas en la Guía de administración de Amazon EMR.
- Al reducir verticalmente el tamaño de un clúster, Amazon EMR copia los datos de las instancias que elimina a las instancias restantes. Asegúrese de que haya suficiente capacidad de almacenamiento para estos datos en las instancias que permanecen en el grupo.
- Amazon EMR intenta retirar HDFS en las instancias del grupo. Antes de reducir el tamaño de un clúster, le recomendamos que minimice las E/S de escritura en HDFS.

- Para tener un control más detallado a la hora de reducir el tamaño de un clúster, puede ver el clúster en la consola y acceder a la pestaña Instancias. Seleccione el ID del grupo de instancias cuyo tamaño desee cambiar. A continuación, use la opción Terminar para las instancias específicas que desee eliminar.

Configuración de los tiempos de espera de aprovisionamiento para controlar la capacidad en Amazon EMR

Cuando usa flotas de instancias, puede configurar los tiempos de espera de aprovisionamiento. Un tiempo de espera de aprovisionamiento indica a Amazon EMR que deje de aprovisionar la capacidad de la instancia si el clúster supera un umbral de tiempo especificado durante el lanzamiento del clúster o las operaciones de escalado del clúster. En los siguientes temas, se explica cómo configurar un tiempo de espera de aprovisionamiento para el lanzamiento del clúster y las operaciones de escalado vertical del clúster.

Temas

- Configurar los tiempos de espera de aprovisionamiento para el lanzamiento del clúster en Amazon EMR
- Personalizar un periodo de tiempo de espera de aprovisionamiento para el cambio de tamaño del clúster en Amazon EMR

Configurar los tiempos de espera de aprovisionamiento para el lanzamiento del clúster en Amazon EMR

Puede definir un periodo de tiempo de espera para aprovisionar instancias de spot para cada flota del clúster. Si Amazon EMR no puede aprovisionar la capacidad de spot, puede elegir terminar el clúster o, en su lugar, aprovisionar capacidad bajo demanda. Si el periodo de tiempo de espera finaliza durante el proceso de cambio de tamaño del clúster, Amazon EMR cancela las solicitudes de spot no aprovisionadas. Las instancias de spot no aprovisionadas no se transfieren a la capacidad bajo demanda.

Siga los pasos que se describen a continuación para personalizar un periodo de tiempo de espera de aprovisionamiento para el lanzamiento del clúster con la consola de Amazon EMR.

Console

Para configurar el tiempo de espera de aprovisionamiento al crear un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En la página Crear clúster, vaya a Configuración del clúster y seleccione Flotas de instancias.
4. En Opción de escalado y aprovisionamiento del clúster, especifique el tamaño de spot de las flotas principales y de tarea.
5. En Configuración del tiempo de espera de spot, seleccione Terminar el clúster tras finalizar el tiempo de espera de spot o Cambiar a bajo demanda tras finalizar el tiempo de espera de spot. A continuación, especifique el periodo de tiempo de espera para el aprovisionamiento de las instancias de spot. El valor predeterminado es de 1 hora.
6. Seleccione cualquier otra opción que se aplique a su clúster.
7. Para lanzar el clúster con el tiempo de espera configurado, seleccione Crear clúster.

AWS CLI

Para especificar un tiempo de espera de aprovisionamiento con el comando **`create-cluster`**

```
aws emr create-cluster \
--release-label emr-5.35.0 \
--service-role EMR_DefaultRole \
--ec2-attributes '{"InstanceProfile":"EMR_EC2_DefaultRole","SubnetIds":["subnet-
XXXXX"]}' \
--instance-fleets
 '[{"InstanceFleetType":"MASTER","TargetOnDemandCapacity":1,"TargetSpotCapacity":0,"LaunchSp
{"OnDemandSpecification":{"AllocationStrategy":"lowest-
price"}},"InstanceTypeConfigs":[{"WeightedCapacity":1,"EbsConfiguration":
{"EbsBlockDeviceConfigs":[{"VolumeSpecification":
{"SizeInGB":32,"VolumeType":"gp2"},"VolumesPerInstance":2}]},"BidPriceAsPercentageOfOnDemand
 - 1"},
{"InstanceFleetType":"CORE","TargetOnDemandCapacity":1,"TargetSpotCapacity":1,"LaunchSpecifi
{"SpotSpecification":
{"TimeoutDurationMinutes":120,"TimeoutAction":"SWITCH_TO_ON_DEMAND"},"OnDemandSpecification"
{"AllocationStrategy":"lowest-price"}},"InstanceTypeConfigs":
[{"WeightedCapacity":1,"EbsConfiguration":{"EbsBlockDeviceConfigs":
```

```
[{"VolumeSpecification":
{"SizeInGB":32,"VolumeType":"gp2"},"VolumesPerInstance":2}]},"BidPriceAsPercentageOfOnDemand
 - 2"}]'
```

Personalizar un periodo de tiempo de espera de aprovisionamiento para el cambio de tamaño del clúster en Amazon EMR

Puede definir un periodo de tiempo de espera para aprovisionar instancias de spot para cada flota del clúster. Si Amazon EMR no puede aprovisionar la capacidad de spot, cancela la solicitud de cambio de tamaño y detiene los intentos de aprovisionar capacidad de spot adicional. Al crear un clúster, puede configurar el tiempo de espera. Para un clúster en ejecución, puede agregar o actualizar un tiempo de espera.

Cuando vence el tiempo de espera, Amazon EMR envía automáticamente los eventos a una transmisión de CloudWatch Amazon Events. Con él CloudWatch, puede crear reglas que coincidan con los eventos de acuerdo con un patrón específico y, a continuación, enrutar los eventos a los objetivos para que tomen medidas. Por ejemplo, puede configurar una regla para enviar una notificación por correo electrónico. Para obtener más información sobre cómo crear reglas, consulte Creación de reglas para eventos de Amazon EMR con CloudWatch. Para obtener más información sobre los diferentes detalles de eventos, consulte Eventos de cambio de estado de la flota de instancias.

Ejemplos de tiempos de espera de aprovisionamiento agotados para cambiar el tamaño del clúster

Especificar un tiempo de espera de aprovisionamiento para el cambio de tamaño con la AWS CLI

En el siguiente ejemplo, se utiliza el comando `create-cluster` para agregar un tiempo de espera de aprovisionamiento para cambiar el tamaño.

```
aws emr create-cluster \
--release-label emr-5.35.0 \
--service-role EMR_DefaultRole \
--ec2-attributes '{"InstanceProfile":"EMR_EC2_DefaultRole","SubnetIds":["subnet-
XXXXX"]}' \
--instance-fleets
 '[{"InstanceFleetType":"MASTER","TargetOnDemandCapacity":1,"TargetSpotCapacity":0,"InstanceTyp
[{"WeightedCapacity":1,"EbsConfiguration":{"EbsBlockDeviceConfigs":
[{"VolumeSpecification":
{"SizeInGB":32,"VolumeType":"gp2"},"VolumesPerInstance":2}]},"BidPriceAsPercentageOfOnDemandPri
```

```
 - 1"},
{"InstanceFleetType":"CORE","TargetOnDemandCapacity":1,"TargetSpotCapacity":1,"LaunchSpecificat
{"SpotSpecification":
{"TimeoutDurationMinutes":120,"TimeoutAction":"SWITCH_TO_ON_DEMAND"},"OnDemandSpecification":
{"AllocationStrategy":"lowest-price"}},"ResizeSpecifications":
{"SpotResizeSpecification":{"TimeoutDurationMinutes":20},"OnDemandResizeSpecification":
{"TimeoutDurationMinutes":25}},"InstanceTypeConfigs":
[{"WeightedCapacity":1,"EbsConfiguration":{"EbsBlockDeviceConfigs":
[{"VolumeSpecification":
{"SizeInGB":32,"VolumeType":"gp2"},"VolumesPerInstance":2}]},"BidPriceAsPercentageOfOnDemandPri
 - 2"}]'
```

En el siguiente ejemplo, se utiliza el comando `modify-instance-fleet` para agregar un tiempo de espera de aprovisionamiento para cambiar el tamaño.

```
aws emr modify-instance-fleet \
--cluster-id j-XXXXXXXXXXXXX \
--instance-fleet '{"InstanceFleetId":"if-XXXXXXXXXXXX","ResizeSpecifications":
{"SpotResizeSpecification":{"TimeoutDurationMinutes":30},"OnDemandResizeSpecification":
{"TimeoutDurationMinutes":60}}}' \
--region us-east-1
```

En el siguiente ejemplo, se utiliza `add-instance-fleet-command` para agregar un tiempo de espera de aprovisionamiento para cambiar el tamaño.

```
aws emr add-instance-fleet \
--cluster-id j-XXXXXXXXXXXXX \
--instance-fleet
 '{"InstanceFleetType":"TASK","TargetOnDemandCapacity":1,"TargetSpotCapacity":0,"InstanceTypeCo
[{"WeightedCapacity":1,"EbsConfiguration":{"EbsBlockDeviceConfigs":
[{"VolumeSpecification":
{"SizeInGB":32,"VolumeType":"gp2"},"VolumesPerInstance":2}]},"BidPriceAsPercentageOfOnDemandPri
{"SpotResizeSpecification":{"TimeoutDurationMinutes":30},"OnDemandResizeSpecification":
{"TimeoutDurationMinutes":35}}}' \
--region us-east-1
```

Especifique un tiempo de espera de aprovisionamiento para cambiar el tamaño y lanzarlo con AWS CLI

En el siguiente ejemplo, se utiliza el comando `create-cluster` para agregar un tiempo de espera de aprovisionamiento para cambiar el tamaño y realizar el lanzamiento.

```
aws emr create-cluster \
--release-label emr-5.35.0 \
--service-role EMR_DefaultRole \
--ec2-attributes '{"InstanceProfile":"EMR_EC2_DefaultRole","SubnetIds":["subnet-
XXXXX"]}' \
--instance-fleets
 '[{"InstanceFleetType":"MASTER","TargetOnDemandCapacity":1,"TargetSpotCapacity":0,"LaunchSpeci
{"OnDemandSpecification":{"AllocationStrategy":"lowest-price"}},"InstanceTypeConfigs":
[{"WeightedCapacity":1,"EbsConfiguration":{"EbsBlockDeviceConfigs":
[{"VolumeSpecification":
{"SizeInGB":32,"VolumeType":"gp2"},"VolumesPerInstance":2}]},"BidPriceAsPercentageOfOnDemandPri
 - 1"},
{"InstanceFleetType":"CORE","TargetOnDemandCapacity":1,"TargetSpotCapacity":1,"LaunchSpecificat
{"SpotSpecification":
{"TimeoutDurationMinutes":120,"TimeoutAction":"SWITCH_TO_ON_DEMAND"},"OnDemandSpecification":
{"AllocationStrategy":"lowest-price"}},"ResizeSpecifications":
{"SpotResizeSpecification":{"TimeoutDurationMinutes":20},"OnDemandResizeSpecification":
{"TimeoutDurationMinutes":25}},"InstanceTypeConfigs":
[{"WeightedCapacity":1,"EbsConfiguration":{"EbsBlockDeviceConfigs":
[{"VolumeSpecification":
{"SizeInGB":32,"VolumeType":"gp2"},"VolumesPerInstance":2}]},"BidPriceAsPercentageOfOnDemandPri
 - 2"}]'
```

Consideraciones para los tiempos de espera de aprovisionamiento para el cambio de tamaño

Al configurar los tiempos de espera de aprovisionamiento del clúster para las flotas de instancias, tenga en cuenta los siguientes comportamientos.

- Puede configurar tiempos de espera de aprovisionado para instancias de spot y bajo demanda. El tiempo de espera mínimo de aprovisionamiento es de 5 minutos. El tiempo de espera máximo de aprovisionamiento es de 7 días.
- Solo puede configurar los tiempos de espera de aprovisionamiento para un clúster de EMR que utilice flotas de instancias. Debe configurar cada núcleo y cada flota de tareas por separado.
- Al crear un clúster, puede configurar los tiempos de espera de aprovisionamiento. Puede agregar un tiempo de espera o actualizar uno existente para un clúster en ejecución.
- Si envía varias operaciones de cambio de tamaño, Amazon EMR realiza un seguimiento de los tiempos de espera de aprovisionamiento para cada operación de cambio de tamaño. Por ejemplo, establezca el tiempo de espera de aprovisionamiento en un clúster en minutos. *60* A continuación, envíe una operación de cambio de tamaño cada vez. *R1 T1* Realice una segunda operación de cambio de tamaño cada *R2* vez. *T2* El tiempo de espera de aprovisionamiento para R1 vence a

las. *T1 + 60 minutes* El tiempo de espera de aprovisionamiento para R2 vence a las. *T2 + 60 minutes*

- Si envía una nueva operación de aumento de tamaño de escalado vertical antes de que se agote el tiempo de espera, Amazon EMR continúa intentando aprovisionar capacidad para su clúster de EMR.

Opciones de reducción vertical de clústeres para clústeres de Amazon EMR

> **Note**
>
> Las opciones de comportamiento de reducción vertical ya no se admiten desde la versión 5.10.0 de Amazon EMR. Gracias a la introducción de la facturación por segundo en Amazon EC2, el comportamiento de reducción predeterminado de los clústeres de Amazon EMR ahora finaliza al finalizar la tarea.

Con las versiones 5.1.0 a 5.9.1 de Amazon EMR, hay dos opciones para el comportamiento de reducción: finalizar en el límite de una hora de la instancia para la facturación de Amazon o finalizar al finalizar la tarea. EC2 A partir de la versión 5.10.0 de Amazon EMR, la configuración de rescisión en el límite de una hora de instancia ha quedado obsoleta debido a la introducción de la facturación por segundo en Amazon. EC2 No recomendamos especificar la terminación en el límite por instancia y hora en las versiones en las que la opción está disponible.

> **Warning**
>
> Si utiliza la AWS CLI para emitir un`EC2InstanceIdsToTerminate`, estas instancias se cancelan inmediatamente, sin tener en cuenta esta configuración y `modify-instance-groups` con independencia del estado de las aplicaciones que se estén ejecutando en ellas. Terminar una instancia de esta forma conlleva el riesgo de pérdida de datos y de un comportamiento del clúster imprevisible.

Si se especifica la terminación al finalizar la tarea, Amazon EMR deniega las listas y las vacía de los nodos antes de terminar las instancias de Amazon. EC2 Con cualquiera de los dos comportamientos especificados, Amazon EMR no termina las instancias de Amazon en EC2 los grupos de instancias principales si esto pudiera provocar daños en el HDFS.

Terminación al completar las tareas

Amazon EMR le permite reducir verticalmente el clúster sin que ello afecte a su carga de trabajo. Amazon EMR intenta desmantelar fácilmente YARN, HDFS y otros demonios de los nodos principales y de tareas durante una operación de redimensionamiento sin perder datos ni interrumpir las tareas. Amazon EMR solo reduce el tamaño de los grupos de instancias si el trabajo asignados a los grupos se ha completado y están inactivos. En el caso de YARN NodeManager Graceful Decommission, puedes ajustar manualmente el tiempo que espera un nodo a su desmantelamiento.

> **Note**
>
> Cuando se produce un desmantelamiento correcto, puede producirse una pérdida de datos. Asegúrese de hacer una copia de seguridad de sus datos.

> **Important**
>
> Es posible que los datos de HDFS se pierdan de forma permanente al sustituir sin problemas una instancia principal en mal estado. Le recomendamos que siempre haga una copia de seguridad de sus datos.

Este tiempo se establece mediante una propiedad de la clasificación de configuración `YARN-site`. Si utiliza la versión 5.12.0 o una posterior de Amazon EMR, especifique la propiedad `YARN.resourcemanager.nodemanager-graceful-decommission-timeout-secs`. Si utiliza versiones anteriores de Amazon EMR, especifique la propiedad `YARN.resourcemanager.decommissioning.timeout`.

Si sigue habiendo contenedores en ejecución en aplicaciones YARN cuando transcurre el tiempo de espera de retirada, se obliga la retirada del nodo y las reprogramaciones de YARN afectan a contenedores de otros nodos. El valor predeterminado es de 3 600 segundos (una hora). Puede definir este tiempo de espera con un valor arbitrariamente alto para obligar a que la reducción con gracia espere más. Para obtener más información, consulte Graceful Decommission of YARN Nodes en la documentación de Apache Hadoop.

Grupos de nodos de tarea

Amazon EMR selecciona de manera inteligente las instancias que no tienen tareas ejecutándose en ningún paso o aplicación y las elimina de un clúster en primer lugar. Si se están utilizando todas las

instancias del clúster, Amazon EMR espera a que las tareas se completen en una instancia antes de eliminarla del clúster. El tiempo de espera predeterminado es de 1 hora. Este valor se puede cambiar con la configuración `YARN.resourcemanager.decommissioning.timeout`. Amazon EMR utiliza de forma dinámica la nueva configuración. Puede establecerlo en un número arbitrariamente grande para garantizar que Amazon EMR no termine ninguna tarea y, al mismo tiempo, reduzca el tamaño del clúster.

Grupos de nodos principales

En los nodos principales, se deben retirar los DataNode daemons de YARN NodeManager y HDFS para reducir el grupo de instancias. Para YARN, la reducción con gracia garantiza que un nodo marcado para retirada solo pase al estado `DECOMMISSIONED` si no hay contenedores o aplicaciones pendientes o incompletos. La retirada finaliza de inmediato si no hay contenedores en ejecución en el nodo al principio de la retirada.

Para HDFS, la reducción con gracia garantiza que la capacidad de destino de HDFS sea lo suficientemente grande como para adaptarse a todos los bloques existentes. Si la capacidad de destino no es lo suficientemente grande, solo una cantidad parcial de las instancias secundarios se retiran de forma que los nodos restantes pueden gestionar los datos actuales residentes en HDFS. Debe garantizar una capacidad de HDFS adicional para permitir una retirada adicional. También debes intentar minimizar la escritura I/O antes de intentar reducir los grupos de instancias. Una escritura excesiva I/O puede retrasar la finalización de la operación de cambio de tamaño.

Otro límite es el factor de replicación predeterminado, `dfs.replication` dentro de `/etc/hadoop/conf/hdfs-site`. Cuando crea un clúster, Amazon EMR configura el valor en función del número de instancias en el clúster: 1 con 1-3 instancias, 2 para clústeres con 4-9 instancias y 3 para clústeres con 10 o más instancias.

> **Warning**
>
> 1. Establecer `dfs.replication` en 1 en clústeres con menos de cuatro nodos puede conllevar la pérdida de datos del HDFS si un solo nodo deja de funcionar. Se recomienda que utilice un clúster con al menos cuatro nodos principales para las cargas de trabajo de producción.
> 2. Amazon EMR no permitirá que los clústeres escalen los nodos principales por debajo de `dfs.replication`. Por ejemplo, si `dfs.replication = 2`, el número mínimo de nodos principales es 2.

3. Cuando utiliza el escalado administrado, el escalado automático o decide cambiar el tamaño del clúster manualmente, se recomienda que establezca `dfs.replication` en 2 o más.

La reducción con gracia no permite reducir los nodos principales por debajo del factor de replicación de HDFS. Esto permite que HDFS cierre archivos debido a la falta de réplicas. Para evitar este límite, reduzca el factor de replicación y reinicie el NameNode daemon.

Configurar el comportamiento de reducción vertical de Amazon EMR

> **Note**
>
> La opción de comportamiento de reducción vertical con terminación de la instancia después de la hora ya no es compatible con la versión 5.10.0 y posteriores de Amazon EMR. Las siguientes opciones de comportamiento de reducción vertical solo aparecen en la consola de Amazon EMR para las versiones de 5.1.0 a 5.9.1.

Puede usar la AWS Management Console API Amazon EMR o la API de Amazon EMR para configurar el comportamiento de reducción al crear un clúster. AWS CLI

Console

Para configurar el comportamiento de reducción vertical con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres y, a continuación, elija Crear clúster.
3. En la sección de Opción de escalado y aprovisionamiento de clústeres, elija Uso de escalado automático personalizado. En Políticas de escalado automático personalizadas, pulse el botón de acción “más” para añadir las políticas de reducción horizontal. Le recomendamos que añada políticas de reducción horizontal y políticas de escalado horizontall. Al agregar solo un conjunto de políticas, Amazon EMR solo realizará un escalado unidireccional y usted tendrá que realizar las demás acciones manualmente.
4. Elija cualquier otra opción que se aplique a su clúster.

5. Para lanzar el clúster, elija Crear clúster.

AWS CLI

Para configurar el comportamiento de reducción de escala con AWS CLI

- Use la opción `--scale-down-behavior` para especificar `TERMINATE_AT_INSTANCE_HOUR` o `TERMINATE_AT_TASK_COMPLETION`.

Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera.

En esta sección se describen los métodos de terminación de un clúster. Para obtener información sobre cómo habilitar la protección de terminación y la terminación automática de clústeres, consulte Control de la finalización de clústeres de Amazon EMR. Puede terminar clústeres en los estados `STARTING`, `RUNNING`o `WAITING`. Un clúster en el estado `WAITING` se debe terminar o se ejecuta de forma indefinida, generando cargos en su cuenta. Puede terminar un clúster que no sea capaz de salir del estado `STARTING` o no pueda completar un paso.

Si está terminando un clúster que tiene definida la protección de terminación, primero debe anular la protección para de poder terminar el clúster. Los clústeres se pueden terminar mediante la consola AWS CLI, la API o mediante programación. `TerminateJobFlows`

Según la configuración del clúster, el clúster puede tardar entre 5 y 20 minutos en finalizar por completo y liberar los recursos asignados, como EC2 las instancias.

> **Note**
>
> No puede reiniciar un clúster terminado, pero puede clonar un clúster terminado para reutilizar su configuración en un clúster nuevo. Para obtener más información, consulte Clonación de un clúster de Amazon EMR utilizando la consola.

> **Important**
>
> Amazon EMR utiliza la función de servicio Amazon EMR y la `AWSServiceRoleForEMRCleanup` función para limpiar los recursos del clúster de su

cuenta que ya no utiliza, como las instancias de Amazon. EC2 Debe incluir acciones para que las políticas de rol eliminen o terminen los recursos. De lo contrario, Amazon EMR no podrá realizar estas acciones de limpieza y podría incurrir en costos por los recursos no utilizados que permanecen en el clúster.

Terminar un clúster con la consola

Puede terminar uno o varios clústeres mediante la consola de Amazon EMR. Los pasos para terminar un clúster en la consola varían en función de si la protección de terminación está activada o desactivada. Para terminar un clúster protegido, primero debe deshabilitar la protección de terminación.

Console

Para finalizar un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. Seleccione Clústeres y, a continuación, elija el clúster que desea terminar.
3. En el menú desplegable Acciones, seleccione Terminar clúster para abrir la ventana Terminar clúster.
4. Cuando se le solicite, seleccione Terminar. Según la configuración del clúster, la terminación puede tardar de 5 a 10 minutos. Para obtener más información sobre cómo terminar clústeres en Amazon EMR, consulte Finalización de un clúster de Amazon EMR en estado de inicio, ejecución o espera..

Terminar un clúster con la AWS CLI

Para finalizar un clúster desprotegido mediante el AWS CLI

Para terminar un clúster desprotegido mediante el AWS CLI, utilice el `terminate-clusters` subcomando con el parámetro --cluster-ids.

- Escribe el siguiente comando para terminar un solo clúster y sustituirlo por tu ID de clúster. *j-3KVXXXXXXX7UG*

```
aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG
```

Para terminar varios clústeres, escribe el siguiente comando y *j-WJ2XXXXXX8EU* reemplaza *j-3KVXXXXXXX7UG* y por tu clúster IDs.

```
aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG j-WJ2XXXXXX8EU
```

Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Para terminar un clúster protegido mediante el AWS CLI

Para terminar un clúster protegido mediante el AWS CLI, primero deshabilite la protección de terminación mediante el `modify-cluster-attributes` subcomando con el `--no-termination-protected` parámetro. A continuación, utilice el subcomando `terminate-clusters` con el parámetro `--cluster-ids` para terminarlo.

1. Escriba el siguiente comando para deshabilitar la protección de terminación y sustitúyalo por *j-3KVTXXXXXX7UG* el ID de su clúster.

   ```
   aws emr modify-cluster-attributes --cluster-id j-3KVTXXXXXX7UG --no-termination-
   protected
   ```

2. Para terminar el clúster, escriba el siguiente comando y *j-3KVXXXXXXX7UG* sustitúyalo por su ID de clúster.

   ```
   aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG
   ```

   Para terminar varios clústeres, escribe el siguiente comando y *j-WJ2XXXXXX8EU* reemplaza *j-3KVXXXXXXX7UG* y por tu clúster IDs.

   ```
   aws emr terminate-clusters --cluster-ids j-3KVXXXXXXX7UG j-WJ2XXXXXX8EU
   ```

   Para obtener más información sobre el uso de los comandos de Amazon EMR en AWS CLI, consulte. https://docs.aws.amazon.com/cli/latest/reference/emr

Terminar un clúster con la API

La `TerminateJobFlows` operación finaliza el procesamiento por pasos, carga los datos de registro de Amazon EC2 a Amazon S3 (si está configurado) y finaliza el clúster de Hadoop. Un clúster también termina automáticamente si establece `KeepJobAliveWhenNoSteps` en `False` en una solicitud `RunJobFlows`.

Puede utilizar esta acción para terminar un solo clúster o una lista de clústeres por clúster. IDs

Para obtener más información sobre los parámetros de entrada exclusivos de`TerminateJobFlows`, consulte TerminateJobFlows. Para obtener más información acerca de los parámetros genéricos en la solicitud, consulte Parámetros de solicitud comunes.

Clonación de un clúster de Amazon EMR utilizando la consola

Puede utilizar la consola de Amazon EMR para clonar un clúster, lo que realiza una copia de la configuración del clúster original para utilizarla como base de un nuevo clúster.

Console

Para clonar un clúster con la consola

1. Inicie sesión en y abra la AWS Management Console consola de Amazon EMR en https://console.aws.amazon.com /emr.
2. En EMR activado, EC2 en el panel de navegación izquierdo, elija Clústeres.
3. Para clonar un clúster desde la lista de clústeres
   a. Utilice las opciones de búsqueda y filtro para encontrar el clúster que desea clonar en la vista de lista.
   b. Seleccione la casilla de verificación a la izquierda de la fila del clúster que desee clonar.
   c. La opción Clonar ahora estará disponible en la parte superior de la vista de lista. Seleccione Clonar para iniciar el proceso de clonación. Si el clúster tiene pasos configurados, seleccione Incluir pasos y Continuar si desea clonar los pasos junto con las demás configuraciones del clúster.
   d. Revise la configuración del nuevo clúster que se ha copiado del clúster clonado. Ajuste la configuración si es necesario. Cuando esté satisfecho con la configuración del nuevo clúster, seleccione Crear clúster para lanzar el nuevo clúster.
4. Para clonar un clúster desde la página de detalles de un clúster

a. Para ir a la página de detalles del clúster que desea clonar, seleccione su ID de clúster en la vista de lista de clústeres.
b. En la parte superior de la página de detalles del clúster, seleccione Clonar el clúster en el menú Acciones para iniciar el proceso de clonación. Si el clúster tiene pasos configurados, seleccione Incluir pasos y Continuar si desea clonar los pasos junto con las demás configuraciones del clúster.
c. Revise la configuración del nuevo clúster que se ha copiado del clúster clonado. Ajuste la configuración si es necesario. Cuando esté satisfecho con la configuración del nuevo clúster, seleccione Crear clúster para lanzar el nuevo clúster.

Automatización de los clústeres de Amazon EMR recurrentes con AWS Data Pipeline

AWS Data Pipeline es un servicio que automatiza el movimiento y la transformación de los datos. Puede utilizarlo para programar la transferencia de datos de entrada en Amazon S3 y para programar el lanzamiento de clústeres para procesar dichos datos. Por ejemplo, considere el caso en que tiene un servidor web que graba registros de tráfico. Si desea ejecutar un clúster semanal para analizar los datos de tráfico, puede utilizarlo AWS Data Pipeline para programar esos clústeres. AWS Data Pipeline es un flujo de trabajo basado en datos, de modo que una tarea (lanzar el clúster) puede depender de otra tarea (mover los datos de entrada a Amazon S3). También tiene la funcionalidad de reintento eficaz.

Para obtener más información AWS Data Pipeline, consulte la Guía para AWS Data Pipeline desarrolladores, especialmente los tutoriales sobre Amazon EMR:

- Tutorial: lanzamiento de un flujo de trabajo de Amazon EMR
- Introducción: procese registros web con AWS Data Pipeline Amazon EMR y Hive
- Tutorial: Importación y exportación de Amazon DynamoDB mediante AWS Data Pipeline

Tutoriales de Amazon EMR

Un clúster de EMR se ejecuta en un ecosistema complejo. El siguiente tutorial cubre casos de uso importantes.

Temas

- Amazon EMR activado EC2 : supervisión mejorada con el CloudWatch uso de métricas y registros personalizados
- Supervise las aplicaciones de Apache Spark en Amazon EMR con Amazon CloudWatch
- Supervise el estado de la solicitud de Amazon EMR con la integración CloudWatch

Amazon EMR activado EC2 : supervisión mejorada con el CloudWatch uso de métricas y registros personalizados

Descripción general

Amazon EMR ofrece capacidades de procesamiento de macrodatos potentes y rentables. Para maximizar el rendimiento y la utilización de los recursos, es esencial una supervisión eficaz. Amazon CloudWatch ofrece una observabilidad integral para los clústeres de EMR, lo que le permite realizar un seguimiento de las métricas y los registros en tiempo real. En este documento se describe cómo:

1. Configure el CloudWatch agente para enviar EMR en los registros a EC2 CloudWatch
2. Agregue Hadoop, YARN y HBase métricas personalizadas mediante clasificaciones
3. Supervise las métricas a través de paneles integrados
4. Realice un seguimiento de los registros del clúster mediante grupos de CloudWatch registros

Requisitos previos y antecedentes

De forma predeterminada, Amazon EMR envía las métricas básicas CloudWatch cada cinco minutos sin coste adicional. Con la versión 7.0+ de EMR, puede implementar el CloudWatch agente para:

- Recopile 34 métricas detalladas adicionales en intervalos de un minuto (se aplican cargos adicionales)
- Recopile métricas de todos los nodos del clúster

- Agregue los datos en el nodo principal antes de enviarlos a CloudWatch
- Acceda a las métricas a través de la pestaña Supervisión o la consola de EMR CloudWatch

EMR 7.1 amplía estas capacidades y le permite configurar el agente para capturar métricas especializadas de Hadoop, YARN y componentes. HBase Para los entornos que utilizan Prometheus, las métricas se pueden reenviar a Amazon Managed Service for Prometheus.

CloudWatch Configuración del agente para los registros

Para capturar los inicios de sesión de EMR CloudWatch, cree un archivo cloudwatch-config.json que defina los archivos de registro que se van a recopilar:

cloudwatch-config.json

```
{
  "agent": {"metrics_collection_interval":60,"logfile":"/var/log/
emr-cluster-metrics/amazon-cloudwatch-agent/amazon-cloudwatch-
agent.log","run_as_user":"****","omit_hostname":true},
  "logs": {
    "logs_collected": {
      "files": {
        "collect_list": [
          {
            "file_path": "/mnt/var/log/hadoop-yarn/hadoop-yarn-resourcemanager-*",
            "log_group_name": "/emr/yarn/resourcemnger",
            "log_stream_name": "{instance_id}",
            "publish_multi_logs" : true
          },
          {
            "file_path": "/var/log/hadoop-hdfs/hadoop-hdfs-namenode-*",
            "log_group_name": "/emr/hdfs/namenode",
            "log_stream_name": "{instance_id}",
            "publish_multi_logs" : true
          }
        ]
      }
    }
  }
}
```

Script de Bootstrap para la configuración del agente CloudWatch

Para aplicar su CloudWatch configuración personalizada a los nodos EMR, cree un script de arranque que reinicie el CloudWatch agente con su configuración. Este script garantiza que el agente se ejecute con sus parámetros específicos de recopilación de registros tras el aprovisionamiento del clúster.

Creación del script de Bootstrap

Cree un archivo denominado cloudwatch-agent-bootstrap.sh con el siguiente contenido:

```
#!/bin/bash
set -xe

EMR_SECONDARY_BA_SCRIPT=$(cat << 'EOF'
while true; do
NODEPROVISIONSTATE=$(sed -n '/localInstance [{]/,/[}]/ {/nodeProvisionCheckinRecord
 [{]/,/[}]/ {/status:/ p}}' /emr/instance-controller/lib/info/job-flow-state.txt | awk
 '{ print $2 }')

if [ "$NODEPROVISIONSTATE" == "SUCCESSFUL" ]; then
sleep 10
echo "Running my post provision bootstrap"
NODETYPE=$(cat /mnt/var/lib/instance-controller/extraInstanceData.json | jq -r
 '.instanceRole' | awk '{print tolower($0)}')

# Copy config file on the instance
sudo aws s3 cp s3://amzn-s3-demo-bucket1>/cloudwatch-config.json /etc/emr-cluster-
metrics/amazon-cloudwatch-agent/conf/emr-amazon-cloudwatch-agent.json

# Stop the current agent
sudo /usr/bin/amazon-cloudwatch-agent-ctl -a stop

# Start the agent with the created config file
sudo /usr/bin/amazon-cloudwatch-agent-ctl -a fetch-config -s -m ec2 -c file:/etc/emr-
cluster-metrics/amazon-cloudwatch-agent/conf/emr-amazon-cloudwatch-agent.json

# Status CW Agent
echo "Status CW Agent"
sudo /usr/bin/amazon-cloudwatch-agent-ctl -m ec2 -a status

exit
fi
```

```
sleep 10
done
EOF
)

echo "${EMR_SECONDARY_BA_SCRIPT}" | tee -a /tmp/emr-secondary-ba.sh
chmod u+x /tmp/emr-secondary-ba.sh
/tmp/emr-secondary-ba.sh > /tmp/emr-secondary-ba.log 2>&1 &
exit 0
```

Nota de configuración importante

> **Important**
>
> Antes de cargar el script, **<amzn-s3-demo-bucket1>** sustitúyalo por el nombre real del bucket de S3 en el que guardaste el archivo cloudwatch-config.json del paso anterior. Esto garantiza que el script de arranque pueda recuperar el archivo de configuración durante la inicialización del clúster.

Este script de arranque hará lo siguiente:

- Espere a que se complete el aprovisionamiento del nodo
- Descarga tu configuración personalizada CloudWatch
- Detenga cualquier CloudWatch agente en ejecución
- Reinicie el agente con su configuración específica
- Registre el estado del agente para solucionar problemas

Clasificaciones métricas personalizadas para Hadoop, YARN y HBase

Además de CloudWatch las métricas predeterminadas, puede mejorar sus capacidades de supervisión configurando métricas personalizadas específicas de la aplicación para los componentes del clúster de EMR. La API de configuración de Amazon EMR proporciona una forma flexible de definir exactamente qué métricas desea recopilar.

Configuración de métricas personalizadas

Puede implementar la recopilación de métricas personalizada de dos maneras:

- Durante la creación de clústeres para nuevos clústeres
- Como reconfiguración de los clústeres existentes a través de la consola EMR

Creación de un archivo de clasificación

El archivo de clasificación define qué métricas de componentes específicos deben recopilarse del clúster. A continuación, se muestra un ejemplo de estructura para recopilar métricas de Hadoop personalizadas:

```
[
  {
    "Classification": "emr-metrics",
    "Configurations": [
      {
        "Classification": "emr-hadoop-hdfs-datanode-metrics",
        "Properties": {
          "Hadoop:service=DataNode,name=DataNodeActivity-*":
 "DatanodeNetworkErrors,TotalReadTime,TotalWriteTime,BytesRead,BytesWritten,RemoteBytesRead,Rem
          "otel.metric.export.interval": "30000"
        }
      },
      {
        "Classification": "emr-hadoop-yarn-nodemanager-metrics",
        "Properties": {
          "Hadoop:service=NodeManager,name=JvmMetrics":
 "MemNonHeapUsedM,MemNonHeapCommittedM,MemNonHeapMaxM,MemHeapUsedM,MemHeapCommittedM,MemHeapMax
          "Hadoop:service=NodeManager,name=NodeManagerMetrics":
 "ContainerCpuUtilization,NodeCpuUtilization,ContainersCompleted,ContainersFailed,ContainersKil
          "otel.metric.export.interval": "20000"
        }
      }
    ],
    "Properties": {}
  }
]
```

Pasos para la implementación

1. Cree un archivo JSON con las clasificaciones métricas que desee.
2. Personalice las métricas en función de sus requisitos de supervisión.
3. Guarde el archivo y cárguelo en su bucket de S3.

4. Consulte este archivo cuando cree un clúster nuevo o reconfigure uno existente.

Prácticas recomendadas

- Recopile únicamente métricas que proporcionen información significativa sobre sus cargas de trabajo.
- Tenga en cuenta el intervalo de recopilación de métricas en función de sus necesidades de supervisión.
- Consulte la AWS documentación para obtener la lista completa de las métricas disponibles para cada componente.
- Agrupe las métricas relacionadas dentro de la misma clasificación para una mejor organización.

Este enfoque le permite centrar la supervisión en las métricas más importantes para sus aplicaciones de EMR específicas, lo que le brinda una mayor visibilidad del rendimiento del clúster.

Implementación de un clúster EMR con integración CloudWatch

Siga estos pasos para crear un clúster de Amazon EMR que envíe automáticamente registros y métricas personalizadas a: CloudWatch

Paso 1: Habilitar el agente CloudWatch

Al crear el clúster de EMR a través de la consola de AWS administración:

1. Navegue hasta la sección Aplicaciones durante la creación del clúster.
2. Seleccione las casillas de verificación de sus aplicaciones principales (Hadoop, Spark, etc.).
3. Desplázate para buscar y seleccionar la opción Amazon CloudWatch Agent.
4. Esto habilita al agente en su clúster, lo cual es esencial para recopilar métricas y registros mejorados.

El CloudWatch agente se instalará en todos los nodos del clúster, lo que le permitirá recopilar las métricas del sistema y de las aplicaciones en los intervalos configurados.

Crear un clúster y mostrar los paquetes disponibles.

> **Note**
>
> El CloudWatch agente está disponible en la versión 7.0 y posteriores de EMR. Es necesario habilitar este componente para la recopilación de métricas y el reenvío de registros personalizados que se describen en esta guía.

Paso 2: Añadir la acción Bootstrap para la recopilación de registros

Para configurar el CloudWatch agente para que recopile y reenvíe archivos de registro específicos a CloudWatch:

1. En el asistente de creación de clústeres de EMR, vaya a la sección Acciones de Bootstrap
2. Haga clic en Añadir acción de arranque
3. Selecciona Acción personalizada en el menú desplegable
4. Proporcione un nombre para la acción de arranque (por ejemplo, Configurar CloudWatch el agente)
5. En el campo Ubicación del script, introduzca la ruta S3 a su script cloudwatch-agent-bootstrap .sh (p. ej., s3://your-bucket-name/cloudwatch-agent-bootstrap.sh)
6. Haga clic en Añadir para guardar la acción de arranque

Esta acción de arranque se ejecutará durante el inicio del clúster, lo que garantizará que CloudWatchagent esté correctamente configurada con sus ajustes personalizados para recopilar y reenviar los archivos de registro especificados en el archivo de configuración.

El agente empezará a recopilar registros automáticamente una vez que se hayan aprovisionado los nodos, lo que proporcionará una visibilidad casi en tiempo real de las operaciones del clúster a través CloudWatch de los registros.

Uso de acciones de arranque.

Paso 3: Configurar la recopilación de métricas personalizadas

Para habilitar la recopilación de HBase métricas personalizadas de Hadoop, YARN o más allá del conjunto predeterminado:

1. En el asistente de creación de clústeres de EMR, vaya a la sección Configuraciones.

2. Haga clic en el botón Editar configuraciones para ampliar las opciones de configuración.
3. Seleccione la opción Cargar JSON desde Amazon S3 en el menú desplegable del método de configuración.
4. Introduzca la ruta URI de S3 a su archivo de clasificación de métricas personalizado (p. ej., s3:// amzn-s3-demo-bucket1/ emr-metrics-classification .json).
5. Haga clic en Cargar para analizar la configuración.
6. Compruebe que la configuración aparece correctamente en la interfaz de la consola.
7. Haga clic en Guardar cambios para aplicar estas configuraciones métricas al clúster.

Este paso indica al CloudWatch agente que recopile las métricas de los componentes específicos definidas en el archivo de clasificación. Las métricas se recopilarán en los intervalos especificados en la configuración y se publicarán allí CloudWatch, donde se podrán visualizar y analizar.

Las métricas personalizadas proporcionan información más detallada sobre las características de rendimiento de su clúster, lo que permite una supervisión y solución de problemas más precisas de sus aplicaciones de EMR.

Anule las configuraciones predeterminadas.

Actualización de la configuración de métricas para los clústeres en ejecución

Puede modificar la configuración de recopilación de métricas de un clúster de EMR existente sin interrumpir las operaciones siguiendo estos pasos:

1. Navegue hasta el clúster de EMR activo en la consola de AWS administración.
2. Seleccione la pestaña Configuraciones en la vista de detalles del clúster.
3. Busque la sección de configuraciones de grupos de instancias.
4. Haga clic en el botón Reconfigurar para modificar la configuración.
5. Elija Cargar JSON desde Amazon S3 o edite directamente la configuración.
6. Introduzca la ubicación actualizada del archivo de clasificación de métricas o realice los cambios en el editor.
7. Aplique los cambios para actualizar el comportamiento de recopilación de métricas.

Esta capacidad de reconfiguración le permite ajustar su enfoque de supervisión a medida que evolucionan los requisitos de carga de trabajo. El CloudWatch agente se adaptará automáticamente

a la nueva configuración y recopilará el conjunto de métricas actualizado sin requerir el reinicio del clúster ni el tiempo de inactividad.

> **Important**
>
> Los cambios de configuración pueden tardar varios minutos en propagarse a todos los nodos del clúster. Siga supervisando sus CloudWatch paneles de control para confirmar que las nuevas métricas se muestran según lo esperado.

Configuraciones de grupos de instancias.

Validar la integración CloudWatch

Tras completar los pasos de configuración, es el momento de comprobar que la configuración de supervisión funciona correctamente:

Paso 1: Implemente su clúster de EMR

1. Revise todos los ajustes de configuración para comprobar su precisión.
2. Asegúrese de que las acciones de arranque y los archivos de clasificación estén referenciados correctamente.
3. Haga clic en Crear clúster para iniciar su entorno EMR.
4. Espere a que el clúster alcance el estado En ejecución (normalmente de 5 a 15 minutos).

Paso 2: Ejecute las aplicaciones de prueba

Envía varias aplicaciones de prueba a Spark para generar métricas significativas:

- Ejecuta un trabajo sencillo de Spark que procese datos de muestra.
- Ejecuta una tarea de análisis que dure más tiempo para observar la utilización de los recursos.
- Pruebe diferentes configuraciones de aplicaciones para comparar las métricas de rendimiento.

Una vez completadas las aplicaciones (o mientras están en ejecución):

- Navega hasta la CloudWatch consola.

- Compruebe los grupos de registros configurados para ver los registros de las aplicaciones.
- Examine los paneles de métricas para observar las métricas específicas de la CPU, la memoria y la aplicación.
- Compruebe que las métricas personalizadas definidas en el archivo de clasificación aparezcan en. CloudWatch

Este proceso de validación confirma que su CloudWatch integración captura correctamente tanto los registros como las métricas, lo que le proporciona una visibilidad completa del rendimiento y el comportamiento de las aplicaciones de su clúster de EMR.

Acceso a los registros de EMR en CloudWatch grupos de registros

Una vez que el clúster de EMR esté en ejecución y el CloudWatch agente esté correctamente configurado, los registros de la aplicación y del sistema estarán disponibles en CloudWatch Registros. Siga estos pasos para acceder a ellos y analizarlos:

Visualización de sus grupos de registros

1. Navegue hasta la CloudWatch consola en AWS Management Console.
2. Seleccione Grupos de registros en el panel de navegación izquierdo.
3. Busque los grupos de registros creados por su configuración, como:
   - /emr/yarn/resourcemngerpara los ResourceManager registros de YARN.
   - /emr/hdfs/namenodepara NameNode registros HDFS.
   - Cualquier grupo de registros adicional especificado en el archivo de configuración.

Cada grupo de registros contiene flujos de registros organizados por ID de instancia, lo que le permite rastrear los registros hasta nodos específicos del clúster.

Trabajando con datos de registro

- Busque datos de registro: utilice CloudWatch Logs Insights para realizar consultas estructuradas en sus grupos de registros.
- Cree métricas: extraiga métricas de los patrones de registro para crear CloudWatch métricas personalizadas.
- Configure alertas: configure las alarmas en función de patrones de error específicos o frecuencias de registro.

- Exportación de registros: descargue los registros para analizarlos o archivarlos sin conexión.

Retención de registros

> **Note**
>
> De forma predeterminada, los registros se conservan durante 30 días. Puede modificar la política de retención de cada grupo de registros para conservar los registros durante períodos más largos si es necesario con fines de cumplimiento o análisis.

CloudWatch Los registros proporcionan una ubicación centralizada para todos los datos de registro de EMR, lo que elimina la necesidad de usar SSH en los nodos individuales del clúster para solucionar problemas o analizar el comportamiento de las aplicaciones.

Visualización de métricas personalizadas en el panel de monitoreo de EMR

Una vez que el clúster de EMR se ejecute con el CloudWatch agente y la configuración de métricas personalizadas, puede supervisar fácilmente estas métricas directamente en la consola de EMR:

Acceder a sus métricas personalizadas

1. Navegue hasta el clúster de EMR en la consola de AWS administración.
2. Seleccione la pestaña Supervisión en la página de detalles del clúster.
3. Localice el menú desplegable de clasificación de métricas del filtro cerca de la parte superior de los paneles de monitoreo.
4. Utilice este filtro para seleccionar categorías de métricas específicas:
   - Elija HDFS para ver NameNode las DataNode métricas.
   - Seleccione YARN para ver ResourceManager y almacenar las métricas.
   - Elija HBasedatos HBase de rendimiento específicos.
   - Seleccione las clasificaciones de métricas personalizadas que haya definido.

El panel se actualizará de forma dinámica para mostrar gráficos de las métricas seleccionadas, que muestran las tendencias de rendimiento a lo largo del tiempo.

Trabajando con visualizaciones de métricas

- Ajustar los intervalos de tiempo: cambie la ventana de tiempo para ver la actividad reciente o las tendencias históricas.
- Compare las métricas: muestre varias métricas relacionadas side-by-side para el análisis de correlación.
- Funciones de zoom: concéntrese en períodos de tiempo específicos en los que aparecen anomalías o patrones.
- Actualice los datos: actualice las visualizaciones con los datos de métricas más recientes prácticamente en tiempo real.

Este enfoque de monitoreo integrado le permite realizar un seguimiento tanto de las métricas de EMR estándar como de las métricas personalizadas en un panel unificado, lo que facilita la identificación de los problemas de rendimiento, las limitaciones de recursos o los cuellos de botella de las aplicaciones sin salir de la consola de EMR.

Clasificación de métricas de filtrado.

Supervise las aplicaciones de Apache Spark en Amazon EMR con Amazon CloudWatch

Para mejorar la eficiencia de una aplicación Spark, es esencial monitorear su rendimiento y comportamiento. En la siguiente entrada de blog, un tutorial demuestra de alguna step-by-step manera cómo publicar métricas detalladas de Spark desde Amazon EMR a. CloudWatch Esto le permite identificar los cuellos de botella y, al mismo tiempo, optimizar la utilización de los recursos: supervise las aplicaciones de Apache Spark en Amazon EMR con Amazon CloudWatch

Para obtener más información CloudWatch, consulta ¿Qué es Amazon CloudWatch? .

Supervise el estado de la solicitud de Amazon EMR con la integración CloudWatch

Cuando se integra CloudWatch con Amazon EMR, puede realizar un seguimiento de los estados críticos de aplicaciones como 2. HiveServer Puede publicar el estado en métricas CloudWatch personalizadas y configurar alertas en caso de falta de disponibilidad del servicio.

En concreto, puede crear un script para monitorizar aplicaciones de Amazon EMR como YARN ResourceManager y HiveServer 2 en un nodo principal. Consulte Publicar y supervisar el estado de una solicitud de Amazon EMR con CloudWatch integración en el Centro de conocimiento de Re:post para obtener más información sobre cómo configurar este caso de uso.

Solución de problemas de clústeres de Amazon EMR

Un clúster EMR se ejecuta en un ecosistema complejo que comprende software de código abierto, código de aplicación personalizado y. Servicios de AWS Si se produce un problema con alguna de estas partes, es posible que el clúster falle o tarde más de lo esperado en completarse. Los temas siguientes le ayudarán a identificar problemas del clúster y a solucionarlos.

Temas

- ¿Qué herramientas hay disponibles para la solución de problemas de un clúster de Amazon EMR?
- Ver y reiniciar Amazon EMR y procesos de aplicaciones (daemons)
- Recopilaciones de errores comunes en Amazon EMR
- Solución de problemas de un clúster de Amazon EMR que ha fallado con un código de error
- Solución de problemas de un clúster de Amazon EMR lento
- Solución de problemas comunes al utilizar Amazon EMR AWS con Lake Formation

Al desarrollar una nueva aplicación de Hadoop, le recomendamos que habilite la depuración y procese un subconjunto pequeño, pero representativo, de sus datos para probar la aplicación. También puede ejecutar la aplicación step-by-step para probar cada paso por separado. Para obtener más información, consulte Configuración del registro y la depuración de un clúster de Amazon EMR y Paso 5: compruebe el clúster de Amazon EMR de forma pormenorizada.

¿Qué herramientas hay disponibles para la solución de problemas de un clúster de Amazon EMR?

Para identificar y corregir los errores del clúster, puede utilizar las herramientas que se describen en esta página. Es posible que tenga que inicializar algunas de las herramientas al lanzar el clúster. De forma predeterminada, hay otras herramientas disponibles para cada clúster.

Temas

- Ver detalles del clúster de EMR
- Ver detalles de errores del clúster de EMR
- Ejecutar scripts y configurar procesos de Amazon EMR
- Ver archivos de registro de

- Supervisar el rendimiento del clúster de EMR

Ver detalles del clúster de EMR

Puede usar la API de EMR AWS Management Console AWS CLI, o la API de EMR para recuperar información detallada sobre un clúster de EMR y la ejecución de tareas. Para obtener más información sobre el uso de AWS Management Console y AWS CLI, consulte. Visualización del estado y los detalles del clúster de Amazon EMR

Panel de detalles de la consola de Amazon EMR

En la lista Clústeres de la consola de Amazon EMR puede ver información de alto nivel sobre el estado de cada clúster de su cuenta y su Región de AWS. En la lista, se muestran todos los clústeres activos y terminados que ha lanzado en los dos últimos meses. En la lista Clusters (Clústeres), puede seleccionar el Name (Nombre) de un clúster para ver los detalles del clúster. Esta información está organizada en distintas categorías para poder consultarla más fácilmente.

La opción Interfaces de usuario de aplicaciones, disponible en la página de detalles del clúster, puede ser especialmente útil para la resolución de problemas. Proporciona el estado de las aplicaciones de YARN y, en algunos casos, como en las aplicaciones de Spark, puede explorar las diferentes métricas y facetas, como trabajos, etapas y ejecutores. Para obtener más información, consulte Visualización del historial de aplicaciones de Amazon EMR. Esta característica solo está disponible para las versiones 5.8.0 y posteriores de Amazon EMR.

Interfaz de línea de comandos de Amazon EMR

Puede encontrar detalles sobre un clúster a partir del `--describe` argumento AWS CLI with the.

API de Amazon EMR

Puede encontrar detalles sobre un clúster desde la API utilizando la acción `DescribeJobFlows`.

Ver detalles de errores del clúster de EMR

Cuando un clúster de EMR termina con un error, `ListClusters` APIs devuelve un código de error `DescribeCluster` y un mensaje de error. En el caso de determinados errores del clúster, la matriz de datos `ErrorDetail` puede ayudarle a solucionar el error.

Para obtener una lista de códigos de error que incluyen datos de `ErrorDetail`, consulte Códigos de error con ErrorDetail información en Amazon EMR.

> **Note**
>
> Mejoramos continuamente nuestros mensajes de error para que reciba la información más reciente y pertinente. No recomendamos analizar el texto desde `ErrorMessage` porque está sujeto a cambios.

Ejecutar scripts y configurar procesos de Amazon EMR

Como parte del proceso de resolución de problemas, puede resultarle útil ejecutar scripts personalizados en el clúster o ver y configurar los procesos del clúster.

Ver y reiniciar los procesos de la aplicación

Puede resultar útil ver los procesos en ejecución en el clúster para diagnosticar posibles problemas. Para detener y reiniciar los procesos del clúster, puede conectarse al nodo maestro del clúster. Para obtener más información, consulte Ver y reiniciar Amazon EMR y procesos de aplicaciones (daemons).

Ejecutar comandos y scripts sin una conexión SSH

Para ejecutar un comando o un script en el clúster como paso, puede usar las herramientas `command-runner.jar` o `script-runner.jar` sin establecer una conexión SSH con el nodo maestro. Para obtener más información, consulte Ejecutar comandos y scripts en un clúster de Amazon EMR.

Ver archivos de registro de

Amazon EMR y Hadoop generan archivos de registro cuando se ejecuta el clúster. Puede acceder a estos archivos de registro desde diversas herramientas, en función de la configuración que haya especificado al lanzar el clúster. Para obtener más información, consulte Configuración del registro y la depuración de un clúster de Amazon EMR.

Archivos de registro en el nodo maestro

Cada clúster publica los archivos de registro en el directoriomnt/var/log//del nodo principal. Estos archivos de registro solo están disponibles mientras se ejecuta el clúster.

Archivos de registro archivados en Amazon S3

Si lanza el clúster y especifica una ruta de registro de Amazon S3, el clúster copia los archivos de registro almacenados en/mnt/var/log/del nodo maestro en Amazon S3 en intervalos de 5 minutos. Esto garantiza que tenga acceso a los archivos de registro incluso después de que el clúster se termine. Dado que los archivos están archivados en intervalos de 5 minutos, los últimos minutos de un clúster terminado de forma repentina podrían no estar disponibles.

Supervisar el rendimiento del clúster de EMR

Amazon EMR proporciona varias herramientas para supervisar el rendimiento del clúster.

Interfaces web de Hadoop

Cada clúster publica una serie de interfaces web en el nodo principal que contienen información sobre el clúster. Puede acceder a estas páginas web mediante un túnel SSH para conectarlas en el nodo principal. Para obtener más información, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

CloudWatch métricas

Cada clúster informa de las métricas a CloudWatch. CloudWatch es un servicio web que realiza un seguimiento de las métricas y que puede utilizar para configurar alarmas en esas métricas. Para obtener más información, consulte Supervisión de las métricas de Amazon EMR con CloudWatch.

Ver y reiniciar Amazon EMR y procesos de aplicaciones (daemons)

Cuando realice la solución de problemas en un clúster, conviene que cree una lista de los procesos en ejecución. Es posible que también desee detener o reiniciar procesos. Por ejemplo, puede reiniciar un proceso después de cambiar una configuración o detectar un problema con un proceso determinado tras analizar los archivos de registro y los mensajes de error.

Hay dos tipos de procesos que se ejecutan en un clúster: los procesos de Amazon EMR (por ejemplo, instance-controller y Log Pusher) y los procesos asociados a las aplicaciones instaladas en el clúster (por ejemplo, y). hadoop-hdfs-namenode hadoop-yarn-resourcemanager

Para trabajar con procesos directamente en un clúster, antes debe conectarse al nodo maestro. Para obtener más información, consulte Conectar a un clúster de Amazon EMR.

Ver procesos en ejecución

El método que utilice para ver los procesos en ejecución en un clúster varía según la versión de Amazon EMR que utilice.

EMR 5.30.0 and 6.0.0 and later

Example : enumerar todos los procesos en ejecución

En el siguiente ejemplo, se utiliza `systemctl` y se especifica `--type` para ver todos los procesos.

```
systemctl --type=service
```

Example : enumerar procesos específicos

En el siguiente ejemplo, se muestra una lista de todos los procesos con nombres que contengan `hadoop`.

```
systemctl --type=service | grep -i hadoop
```

Ejemplo de salida:

```
 hadoop-hdfs-namenode.service           loaded active running Hadoop namenode
 hadoop-httpfs.service                  loaded active running Hadoop httpfs
 hadoop-kms.service                     loaded active running Hadoop kms
 hadoop-mapreduce-historyserver.service loaded active running Hadoop historyserver
 hadoop-state-pusher.service            loaded active running Daemon process that
 processes and serves EMR metrics data.
 hadoop-yarn-proxyserver.service        loaded active running Hadoop proxyserver
 hadoop-yarn-resourcemanager.service    loaded active running Hadoop resourcemanager
 hadoop-yarn-timelineserver.service     loaded active running Hadoop timelineserver
```

Example : ver un informe de estado detallado de un proceso específico

En el siguiente ejemplo, se muestra un informe de estado detallado del servicio `hadoop-hdfs-namenode`.

```
sudo systemctl status hadoop-hdfs-namenode
```

Ejemplo de salida:

```
hadoop-hdfs-namenode.service - Hadoop namenode
   Loaded: loaded (/etc/systemd/system/hadoop-hdfs-namenode.service; enabled; vendor
 preset: disabled)
   Active: active (running) since Wed 2021-08-18 21:01:46 UTC; 26min ago
 Main PID: 9733 (java)
    Tasks: 0
   Memory: 1.1M
   CGroup: /system.slice/hadoop-hdfs-namenode.service
           # 9733 /etc/alternatives/jre/bin/java -Dproc_namenode -Xmx1843m -server -
XX:OnOutOfMemoryError=kill -9 %p ...

Aug 18 21:01:37 ip-172-31-20-123 systemd[1]: Starting Hadoop namenode...
Aug 18 21:01:37 ip-172-31-20-123 su[9715]: (to hdfs) root on none
Aug 18 21:01:37 ip-172-31-20-123 hadoop-hdfs-namenode[9683]: starting namenode,
 logging to /var/log/hadoop-hdfs/ha...out
Aug 18 21:01:46 ip-172-31-20-123 hadoop-hdfs-namenode[9683]: Started Hadoop
 namenode:[  OK  ]
Aug 18 21:01:46 ip-172-31-20-123 systemd[1]: Started Hadoop namenode.
Hint: Some lines were ellipsized, use -l to show in full.
```

EMR 4.x - 5.29.0

Example : enumerar todos los procesos en ejecución

En el ejemplo siguiente, se muestra una lista de todos los procesos en ejecución.

```
initctl list
```

EMR 2.x - 3.x

Example : enumerar todos los procesos en ejecución

En el ejemplo siguiente, se muestra una lista de todos los procesos en ejecución.

```
ls /etc/init.d/
```

Detener y reiniciar procesos

Después de determinar qué procesos se están ejecutando, puede detenerlos y, a continuación, reiniciarlos si es necesario.

EMR 5.30.0 and 6.0.0 and later

Example : detener un proceso

En el siguiente ejemplo, se detiene el proceso `hadoop-hdfs-namenode`.

```
sudo systemctl stop hadoop-hdfs-namenode
```

Puede consultar el `status` para comprobar que el proceso se ha detenido.

```
sudo systemctl status hadoop-hdfs-namenode
```

Ejemplo de salida:

```
hadoop-hdfs-namenode.service - Hadoop namenode
  Loaded: loaded (/etc/systemd/system/hadoop-hdfs-namenode.service; enabled; vendor
 preset: disabled)
  Active: failed (Result: exit-code) since Wed 2021-08-18 21:37:50 UTC; 8s ago
Main PID: 9733 (code=exited, status=143)
```

Example : iniciar un proceso

En el siguiente ejemplo, se inicia el proceso `hadoop-hdfs-namenode`.

```
sudo systemctl start hadoop-hdfs-namenode
```

Puede consultar el estado para comprobar que el proceso se está ejecutando.

```
sudo systemctl status hadoop-hdfs-namenode
```

Ejemplo de salida:

```
hadoop-hdfs-namenode.service - Hadoop namenode
   Loaded: loaded (/etc/systemd/system/hadoop-hdfs-namenode.service; enabled; vendor
 preset: disabled)
   Active: active (running) since Wed 2021-08-18 21:38:24 UTC; 2s ago
  Process: 13748 ExecStart=/etc/init.d/hadoop-hdfs-namenode start (code=exited,
 status=0/SUCCESS)
 Main PID: 13800 (java)
    Tasks: 0
   Memory: 1.1M
```

```
  CGroup: /system.slice/hadoop-hdfs-namenode.service
          # 13800 /etc/alternatives/jre/bin/java -Dproc_namenode -Xmx1843m -server
 -XX:OnOutOfMemoryError=kill -9 %p...
```

EMR 4.x - 5.29.0

Example : detener un proceso en ejecución

En el siguiente ejemplo, se detiene el servicio `hadoop-hdfs-namenode`.

```
sudo stop hadoop-hdfs-namenode
```

Example : reiniciar un proceso detenido

En el siguiente ejemplo, se reinicia el servicio `hadoop-hdfs-namenode`. Debe usar el comando `start` y no `restart`.

```
sudo start hadoop-hdfs-namenode
```

Example : comprobar el estado del proceso

A continuación, se obtiene el estado de `hadoop-hdfs-namenode`. Puede utilizar el comando `status` para comprobar que el proceso se ha detenido o se ha iniciado.

```
sudo status hadoop-hdfs-namenode
```

EMR 2.x - 3.x

Example : detener el proceso de una aplicación

En el siguiente ejemplo, se detiene el servicio `hadoop-hdfs-namenode`, que está asociado a la versión de Amazon EMR instalada en el clúster.

```
sudo /etc/init.d/hadoop-hdfs-namenode stop
```

Example : reiniciar el proceso de una aplicación

El siguiente comando de ejemplo reinicia el proceso `hadoop-hdfs-namenode`:

```
sudo /etc/init.d/hadoop-hdfs-namenode start
```

Example : detener un proceso de Amazon EMR

En el siguiente ejemplo, se detiene un proceso, como instance-controller, que no está asociado a la versión de Amazon EMR del clúster.

```
sudo /sbin/stop instance-controller
```

Example : reiniciar un proceso de Amazon EMR

En el siguiente ejemplo, se reinicia un proceso, como instance-controller, que no está asociado a la versión de Amazon EMR del clúster.

```
sudo /sbin/start instance-controller
```

> **Note**
>
> Los comandos `/sbin/start, stop` y `restart` son symlinks a `/sbin/intictl`. Para obtener más información acerca de `initctl`, consulte la página del manual de initctl escribiendo `man initctl` en el símbolo del sistema.

Recopilaciones de errores comunes en Amazon EMR

A veces, los clústeres fallan o procesan los datos con lentitud. En las secciones siguientes se enumeran problemas comunes de clústeres. Entre los errores se incluyen los errores de arranque y los errores de validación, con sugerencias sobre cómo resolverlos.

Temas

- Códigos de error con ErrorDetail información en Amazon EMR
- Errores de recursos durante las operaciones del clústeres de Amazon EMR
- Errores de entrada y salida de clúster durante la operaciones de Amazon EMR
- Errores de permisos durante las operaciones del clúster de Amazon EMR
- Errores de clúster de Hive
- Errores de la VPC durante las operaciones del clúster de Amazon EMR
- Errores del clúster de Amazon EMR de streaming
- Amazon EMR: errores de clúster JAR personalizados

- Errores de Amazon EMR AWS GovCloud (EE. UU. - Oeste)
- Buscar un clúster que falta

Códigos de error con ErrorDetail información en Amazon EMR

Cuando un clúster de EMR termina con un error, `ListClusters` APIs devuelve un código de error `DescribeCluster` y un mensaje de error. En el caso de algunos errores de clústeres, la matriz de datos `ErrorDetail` puede ayudarle a solucionar el error.

Los errores que incluyen una matriz `ErrorDetail` proporcionan los siguientes detalles:

`ErrorCode`

Un código de error único que se puede utilizar para el acceso mediante programación.

`ErrorData`

Una lista de identificadores en pares clave-valor que puede utilizar para la programación o la búsqueda manual. Para obtener una descripción de los valores `ErrorData` que incluye un código de error, consulte la página de resolución de problemas del código de error.

`ErrorMessage`

Descripción del error con un enlace a más información en la documentación de Amazon EMR.

> **Note**
>
> No recomendamos analizar el texto desde `ErrorMessage` porque está sujeto a cambios.

Códigos de error por categoría

- Códigos de errores de arranque en Amazon EMR
- Códigos de errores internos de Amazon EMR
- Códigos de errores de validación en Amazon EMR

Códigos de errores de arranque en Amazon EMR

En las siguientes secciones, se proporciona información sobre la resolución de problemas relacionados con los códigos de errores de arranque.

Temas

- BOOTSTRAP_FAILURE_PRIMARY_WITH_NON_ZERO_CODE
- BOOTSTRAP_FAILURE_BA_DOWNLOAD_FAILED_PRIMARY
- BOOTSTRAP_FAILURE_FILE_NOT_FOUND_PRIMARY
- BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_PRIMARY
- BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_WORKER
- BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_PRIMARY
- BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_WORKER

BOOTSTRAP_FAILURE_PRIMARY_WITH_NON_ZERO_CODE

Descripción general

Cuando un clúster termina con un error `BOOTSTRAP_FAILURE_PRIMARY_WITH_NON_ZERO_CODE`, se produce un error en una acción de arranque en la instancia principal. Para obtener más información sobre las acciones de arranque, consulte Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR.

Resolución

Para resolver este error, revise los detalles que se muestran en el error de la API, modifique el script de acción de arranque y cree un clúster nuevo con la acción de arranque actualizada.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`primary-instance-id`

El ID de la instancia principal en la que se produjo un error en la acción de arranque.

`bootstrap-action`

El número ordinal de la acción de arranque que falló. Un script con un valor de `bootstrap-action` de `1` es la primera acción de arranque que se ejecuta en la instancia.

`return-code`

El código de retorno de la acción de arranque que falló.

amazon-s3-path

La ubicación en Amazon S3 de la acción de arranque que falló.

public-doc

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir la causa raíz del error de la acción de arranque. A continuación, lance un clúster nuevo.

1. Revise los archivos de registro de acciones de arranque en Amazon S3 para identificar la causa raíz del error. Para obtener más información sobre cómo ver los registros de Amazon EMR, consulte Visualización de los archivos de registro de Amazon EMR.
2. Si activó los registros del clúster al crear la instancia, consulte el registro `stdout` para obtener más información. Puede encontrar el registro `stdout` de la acción de arranque en esta ubicación de Amazon S3:

   ```
   s3://amzn-s3-demo-bucket/logs/Your_Cluster_Id/node/Primary_Instance_Id/bootstrap-actions/Failed_Bootstrap_Action_Number/stdout.gz
   ```

   Para obtener más información sobre registros del clúster, consulte Configuración del registro y la depuración de un clúster de Amazon EMR.
3. Para determinar el error de la acción de arranque, revise las excepciones en los registros `stdout` y el valor `return-code` en `ErrorData`.
4. Utilice los resultados del paso anterior para revisar la acción de arranque de forma que evite las excepciones o pueda gestionarlas correctamente cuando se produzcan.
5. Lance un clúster nuevo con la acción de arranque actualizada.

BOOTSTRAP_FAILURE_BA_DOWNLOAD_FAILED_PRIMARY

Descripción general

Un clúster termina con el error `BOOTSTRAP_FAILURE_BA_DOWNLOAD_FAILED_PRIMARY` cuando la instancia principal no puede descargar un script de acción de arranque desde la ubicación de Amazon S3 que especifique. Las posibles causas son las siguientes:

- El archivo del script de acción de arranque no se encuentra en la ubicación de Amazon S3.
- El rol de servicio de EC2 las instancias de Amazon del clúster (también denominado perfil de EC2 instancia para Amazon EMR) no tiene permisos para acceder al bucket de Amazon S3 en el que reside el script de acción de arranque. Para obtener más información acerca de los roles de servicio, consulte Función de servicio para EC2 instancias de clúster (perfil de EC2 instancia).

Para obtener más información sobre las acciones de arranque, consulte Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR.

Resolución

Para resolver este error, asegúrese de que la instancia principal tenga el acceso adecuado al script de acción de arranque.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`primary-instance-id`

El ID de la instancia principal en la que se produjo un error en la acción de arranque.

`bootstrap-action`

El número ordinal de la acción de arranque que falló. Un script con un valor de `bootstrap-action` de `1` es la primera acción de arranque que se ejecuta en la instancia.

`amazon-s3-path`

La ubicación en Amazon S3 de la acción de arranque que falló.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir la causa raíz del error de la acción de arranque. A continuación, lance un clúster nuevo.

Pasos para la solución de problemas

1. Utilice el valor `amazon-s3-path` de la matriz `ErrorData` para buscar el script de acción de arranque correspondiente en Amazon S3.
2. Si activó los registros del clúster al crear la instancia, consulte el registro `stdout` para obtener más información. Puede encontrar el registro `stdout` de la acción de arranque en esta ubicación de Amazon S3:

   ```
   s3://amzn-s3-demo-bucket/logs/Your_Cluster_Id/node/Primary_Instance_Id/bootstrap-actions/Failed_Bootstrap_Action_Number/stdout.gz
   ```

   Para obtener más información sobre registros del clúster, consulte Configuración del registro y la depuración de un clúster de Amazon EMR.
3. Para determinar el error de la acción de arranque, revise las excepciones en los registros `stdout` y el valor `return-code` en `ErrorData`.
4. Utilice los resultados del paso anterior para revisar la acción de arranque de forma que evite las excepciones o pueda gestionarlas correctamente cuando se produzcan.
5. Lance un clúster nuevo con la acción de arranque actualizada.

BOOTSTRAP_FAILURE_FILE_NOT_FOUND_PRIMARY

Descripción general

El error `BOOTSTRAP_FAILURE_FILE_NOT_FOUND_PRIMARY` indica que la instancia principal no encuentra el script de acción de arranque que la instancia acaba de descargar del bucket de Amazon S3 especificado.

Resolución

Para resolver este error, confirme que la instancia principal tenga el acceso adecuado al script de acción de arranque.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`primary-instance-id`

El ID de la instancia principal en la que se produjo un error en la acción de arranque.

`bootstrap-action`

El número ordinal de la acción de arranque que falló. Un script con un valor de `bootstrap-action` de `1` es la primera acción de arranque que se ejecuta en la instancia.

`amazon-s3-path`

La ubicación en Amazon S3 de la acción de arranque que falló.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir la causa raíz del error de la acción de arranque. A continuación, lance un clúster nuevo.

1. Para encontrar el script de acción de arranque correspondiente en Amazon S3, utilice el valor `amazon-s3-path` de la matriz `ErrorData`.
2. Revise los archivos de registro de acciones de arranque en Amazon S3 para identificar la causa raíz del error. Para obtener más información sobre cómo ver los registros de Amazon EMR, consulte Visualización de los archivos de registro de Amazon EMR.

   > **Note**
   >
   > Si no activó los registros del clúster, debe crear un clúster nuevo con las mismas configuraciones y acciones de arranque. Para asegurarse de que los registros del clúster estén activados, consulte Configuración del registro y la depuración de un clúster de Amazon EMR.

3. Revise el registro `stdout` de las acciones de arranque y confirme que no haya procesos personalizados que eliminen los archivos de la carpeta `/emr/instance-controller/lib/bootstrap-actions` de sus instancias principales. Puede encontrar el registro `stdout` de la acción de arranque en esta ubicación de Amazon S3:

```
s3://amzn-s3-demo-bucket/logs/Your_Cluster_Id/node/Primary_Instance_Id/bootstrap-
actions/Failed_Bootstrap_Action_Number/stdout.gz
```

4. Lance un clúster nuevo con la acción de arranque actualizada.

BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_PRIMARY

Descripción general

El `BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_PRIMARY` error indica que la instancia principal no tiene suficiente espacio en disco al instalar el software necesario.

Resolución

Para resolver este error, confirme que la instancia principal tiene suficiente espacio en disco en el volumen raíz.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`primary-instance-id`

El ID de la instancia principal con espacio en disco insuficiente.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

1. Revise las prácticas recomendadas para el volumen de dispositivos raíz de EBS de su clúster. Consulte Personalización del volumen de dispositivo raíz de Amazon EBS en la Guía de administración de Amazon EMR.
2. Lance un clúster nuevo con un tamaño de volumen de dispositivo raíz de EBS mayor.

BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_WORKER

Descripción general

El `BOOTSTRAP_FAILURE_INSUFFICIENT_DISK_SPACE_WORKER` error indica que una o más instancias de trabajo no tienen suficiente espacio en disco al instalar el software necesario.

Resolución

Para resolver este error, confirme que las instancias de trabajo tienen suficiente espacio en disco en el volumen raíz.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`worker-instance-ids`

La IDs de las instancias de trabajo con espacio en disco insuficiente.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

1. Revise las prácticas recomendadas para el volumen de dispositivos raíz de EBS de su clúster. Consulte Personalización del volumen de dispositivo raíz de Amazon EBS en la Guía de administración de Amazon EMR.
2. Lance un clúster nuevo con un tamaño de volumen de dispositivo raíz de EBS mayor.

BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_PRIMARY

Descripción general

El `BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_PRIMARY` error indica que la instancia principal no puede establecer una conexión con el Hive Metastore externo configurado.

Resolución

Para resolver este error, confirme que su Hive Metastore externo esté configurado correctamente y que la instancia principal pueda conectarse a él.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`primary-instance-id`

El ID de la instancia principal no puede establecer una conexión con el Hive Metastore externo configurado.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

1. Consulte las prácticas recomendadas para configurar un metaalmacén externo para Hive. Consulte Configurar un metabastore externo para Hive.
2. Lance un clúster nuevo con la configuración del clúster actualizada.

BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_WORKER

Descripción general

El `BOOTSTRAP_FAILURE_HIVE_METASTORE_CONNECTION_ERROR_WORKER` error indica que una o más instancias de trabajo no pueden establecer una conexión con el Hive Metastore externo configurado.

Resolución

Para resolver este error, confirme que su Hive Metastore externo esté configurado correctamente y que las instancias de trabajo puedan conectarse a él.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información,

consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`worker-instance-ids`

La IDs de las instancias de trabajo no pudo establecer una conexión con el Hive Metastore externo configurado.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

1. Consulte las prácticas recomendadas para configurar un metaalmacén externo para Hive. Consulte Configurar un metabastore externo para Hive.
2. Lance un clúster nuevo con la configuración del clúster actualizada.

Códigos de errores internos de Amazon EMR

En las siguientes secciones, se proporciona información sobre la resolución de problemas relacionados con los códigos de errores internos, incluidos los códigos para capacidad insuficiente o sin capacidad.

Temas

- INTERNAL_ERROR_ _INSUFFICIENT_CAPACITY_AZ EC2
- INTERNAL_ERROR_SPOT_PRICE_INCREASE_PRIMARY
- INTERNAL_ERROR_SPOT_NO_CAPACITY_PRIMARY

INTERNAL_ERROR_ _INSUFFICIENT_CAPACITY_AZ EC2

Descripción general

Un clúster termina con un `INTERNAL_ERROR_EC2_INSUFFICIENT_CAPACITY_AZ` error cuando la zona de disponibilidad seleccionada no tiene capacidad suficiente para cumplir con tu solicitud de tipo de EC2 instancia de Amazon. La subred que seleccione para un clúster determina la zona de disponibilidad. Para obtener más información sobre las subredes para Amazon EMR, consulte Configuración de redes en una VPC para Amazon EMR.

Resolución

Para resolver este error, modifique las configuraciones del tipo de instancia y cree un clúster nuevo con la solicitud actualizada.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`instance-type`

El tipo de instancia que no tiene capacidad.

`availability-zone`

La zona de disponibilidad en la que se resuelve la subred.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir la causa raíz del error de configuración del clúster:

- Revise las prácticas recomendadas para la configuración del clúster. Consulte Configuración de tipos de instancias de clúster de Amazon EMR y prácticas recomendadas para instancias de Spot en la Guía de administración de Amazon EMR.
- Solucione los problemas de lanzamiento y revise la configuración. Consulta Solución de problemas de lanzamiento de instancias en la Guía del EC2 usuario de Amazon.
- Lance un clúster nuevo con la configuración del clúster actualizada.

INTERNAL_ERROR_SPOT_PRICE_INCREASE_PRIMARY

Descripción general

Un clúster termina con un error `INTERNAL_ERROR_SPOT_PRICE_INCREASE_PRIMARY` cuando Amazon EMR no puede gestionar la solicitud de instancia de spot para el nodo principal porque las instancias no están disponibles por el precio máximo de spot o uno inferior. Para obtener más información, consulte Spot Instances en la Guía del EC2 usuario de Amazon.

Resolución

Para resolver este error, especifique los tipos de instancia para su clúster que estén dentro de su precio objetivo o aumente el límite de precio para el mismo tipo de instancia.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`primary-instance-id`

El ID de la instancia principal del clúster que falló.

`instance-type`

El tipo de instancia que no tiene capacidad.

`availability-zone`

La zona de disponibilidad en la que reside la subred.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para solucionar los problemas de su estrategia de configuración del clúster y, a continuación, lance un clúster nuevo:

1. Revise las prácticas recomendadas para Amazon EC2 Spot Instances y revise su estrategia de configuración de clústeres. Para obtener más información, consulte las prácticas recomendadas para EC2 Spot en la Guía del EC2 usuario de Amazon yConfiguración de tipos de instancias de clúster de Amazon EMR y prácticas recomendadas para instancias de Spot.
2. Modifique las configuraciones del tipo de instancia o la zona de disponibilidad y cree un clúster nuevo con la solicitud actualizada.
3. Si el problema persiste, use la capacidad bajo demanda para la instancia principal.

INTERNAL_ERROR_SPOT_NO_CAPACITY_PRIMARY

Descripción general

Un clúster termina con un error `INTERNAL_ERROR_SPOT_NO_CAPACITY_PRIMARY` cuando no hay suficiente capacidad para gestionar una solicitud de instancia de spot para su nodo principal. Para obtener más información, consulte Spot Instances en la Guía del EC2 usuario de Amazon.

Resolución

Para resolver este error, especifique los tipos de instancia para su clúster que estén dentro de su precio objetivo o aumente el límite de precio para el mismo tipo de instancia.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`primary-instance-id`

El ID de la instancia principal del clúster que falló.

`instance-type`

El tipo de instancia que no tiene capacidad.

`availability-zone`

La zona de disponibilidad en la que se resuelve la subred.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para solucionar los problemas de su estrategia de configuración del clúster y, a continuación, lance un clúster nuevo:

1. Revise las prácticas recomendadas para Amazon EC2 Spot Instances y revise su estrategia de configuración de clústeres. Para obtener más información, consulte las prácticas recomendadas para EC2 Spot en la Guía del EC2 usuario de Amazon yConfiguración de tipos de instancias de clúster de Amazon EMR y prácticas recomendadas para instancias de Spot.

2. Modifique las configuraciones del tipo de instancia y cree un clúster nuevo con la solicitud actualizada.
3. Si el problema persiste, use la capacidad bajo demanda para la instancia principal.

Códigos de errores de validación en Amazon EMR

En las siguientes secciones, se proporciona información sobre la resolución de problemas relacionados con los códigos de errores de validación.

Temas

- VALIDATION_ERROR_SUBNET_NOT_FROM_ONE_VPC
- VALIDATION_ERROR_SECURITY_GROUP_NOT_FROM_ONE_VPC
- VALIDATION_ERROR_INVALID_SSH_KEY_NAME
- VALIDATION_ERROR_INSTANCE_TYPE_NOT_SUPPORTED

VALIDATION_ERROR_SUBNET_NOT_FROM_ONE_VPC

Descripción general

Cuando su clúster y las subredes a las que hace referencia pertenecen a nubes privadas virtuales diferentes (VPCs), el clúster termina con un `VALIDATION_ERROR_SUBNET_NOT_FROM_ONE_VPC` error. Puede lanzar clústeres con Amazon EMR con la configuración de las flotas de instancias en las subredes de una VPC. Para obtener más información sobre las flotas de instancias, consulte Planificación y configuración de flotas de instancias para su clúster de Amazon EMR en la Guía de administración de Amazon EMR.

Resolución

Para resolver este error, utilice subredes que pertenezcan a la misma VPC que el clúster.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

vpc

Para cada par subred:VPC, el ID de la VPC a la que pertenece la subred.

`subnet`

Para cada par subred:VPC, el ID de la subred.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir el error:

1. Revise las subredes IDs que aparecen en la `ErrorData` matriz y confirme que pertenecen a la VPC en la que desea lanzar el clúster de EMR.
2. Modifique las configuraciones de la subred. Puede usar uno de los siguientes métodos para buscar todas las subredes públicas y privadas disponibles en una VPC.
   - Vaya a la consola de Amazon VPC. Elija Subredes y enumere todas las subredes que residen en ellas para su clúster. Región de AWS Para buscar solo subredes públicas o privadas, aplique el filtro de asignación automática de direcciones públicas. IPv4 Para buscar y seleccionar subredes en la VPC que utiliza el clúster, utilice la opción Filtrar por VPC. Para obtener más información sobre cómo crear subredes, consulte Creación de una subred en la Guía del usuario de Amazon Virtual Private Cloud.
   - Use el AWS CLI para buscar todas las subredes públicas y privadas disponibles en la VPC que usa su clúster. Para obtener más información, consulte la API describe-subnets. Para crear subredes en una VPC, consulta la API create-subnet.
3. Lance un clúster nuevo con subredes desde la misma VPC que el clúster.

VALIDATION_ERROR_SECURITY_GROUP_NOT_FROM_ONE_VPC

Descripción general

Cuando el clúster y los grupos de seguridad que le asignas pertenecen a nubes privadas virtuales diferentes (VPCs), el clúster finaliza con un error. `VALIDATION_ERROR_SECURITY_GROUP_NOT_FROM_ONE_VPC` Para obtener más información sobre los grupos de seguridad, consulte Especificación de los grupos de seguridad adicionales administrados por Amazon EMR y Control del tráfico de red con grupos de seguridad para su clúster de Amazon EMR.

Resolución

Para resolver este error, utilice grupos de seguridad que pertenezcan a la misma VPC que el clúster.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`vpc`

Para cada par security-group:VPC, el ID de la VPC a la que pertenece el grupo de seguridad.

`security-group`

Para cada par security-group:VPC, el ID del grupo de seguridad.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir el error:

1. Revise los grupos de seguridad IDs que aparecen en la `ErrorData` matriz y confirme que pertenecen a la VPC en la que desea lanzar el clúster de EMR.
2. Vaya a la consola de Amazon VPC. Elija Grupos de seguridad para enumerar todos los grupos de seguridad de la región que seleccione. Busque los grupos de seguridad de la misma VPC que el clúster y, a continuación, modifique la configuración del grupo de seguridad.
3. Lance un clúster nuevo con grupos de seguridad desde la misma VPC que el clúster.

VALIDATION_ERROR_INVALID_SSH_KEY_NAME

Descripción general

Un clúster termina con un `VALIDATION_ERROR_INVALID_SSH_KEY_NAME` error cuando utilizas un par de EC2 claves de Amazon que no es válido para usar SSH en la instancia principal. Es posible que el nombre del par de claves sea incorrecto o que el par de claves no exista en la solicitud Región de AWS. Para obtener más información sobre los pares de claves, consulta los pares de EC2 claves de Amazon y las instancias de Linux en la Guía del EC2 usuario de Amazon.

Resolución

Para resolver este error, cree un clúster nuevo con un nombre de par de claves SSH válido.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`ssh-key`

El nombre del par de claves SSH que proporcionó al crear el clúster.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir el error:

1. Compruebe el *`keypair`* archivo.pem y confirme que coincide con el nombre de la clave SSH que ve en la consola de Amazon EMR.
2. Ve a la EC2 consola de Amazon. Compruebe que el nombre de la clave SSH que utilizó esté disponible en el clúster Región de AWS que utiliza. La encontrarás Región de AWS junto a tu ID de cuenta en la parte superior de AWS Management Console.
3. Lance un clúster nuevo con un nombre de clave SSH válido.

VALIDATION_ERROR_INSTANCE_TYPE_NOT_SUPPORTED

Descripción general

Un clúster termina con un error `VALIDATION_ERROR_INSTANCE_TYPE_NOT_SUPPORTED` cuando la Región de AWS y las zonas de disponibilidad del clúster no admiten el tipo de instancia especificado para uno o más grupos de instancias. Amazon EMR puede admitir un tipo de instancia en una zona de disponibilidad de una región, pero no en otra. La subred que seleccione para un clúster determina la zona de disponibilidad dentro de la región. Para ver una lista de tipos de instancia y regiones compatibles con Amazon EMR, consulte Tipos de instancias admitidas con Amazon EMR.

Resolución

Para resolver este error, especifique los tipos de instancia para el clúster que Amazon EMR admite en la región y la zona de disponibilidad en la que solicita el clúster.

Para solucionar los problemas del clúster de EMR defectuoso, consulte `ErrorDetail` la información devuelta por y. `DescribeCluster` `ListClusters` APIs Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR. La matriz `ErrorData` de `ErrorDetail` devuelve la siguiente información para este código de error:

`instance-types`

La lista de tipos de instancia no compatibles.

`availability-zones`

La lista de zonas de disponibilidad en las que se resuelve la subred.

`public-doc`

La URL pública de la documentación del código de error.

Pasos que completar

Siga estos pasos para identificar y corregir el error:

1. Úselo AWS CLI para recuperar los tipos de instancias disponibles en una zona de disponibilidad. Para ello, puede usar el `ec2 describe-instance-type-offerings` comando para filtrar los tipos de instancias disponibles por ubicación (Región de AWS o zona de disponibilidad). Por ejemplo, el siguiente comando devuelve los tipos de instancias que se ofrecen en la zona de disponibilidad especificada, *`us-east-2a`*.

   ```
   aws ec2 describe-instance-type-offerings --location-type "availability-
   zone" --filters Name=location,Values=us-east-2a --region us-east-2 --query
    "InstanceTypeOfferings[*].[InstanceType]" --output text | sort
   ```

   Para obtener más información sobre cómo descubrir los tipos de instancias disponibles, consulta Buscar un tipo de EC2 instancia de Amazon.
2. Tras determinar los tipos de instancia que están disponibles en la misma región y zona de disponibilidad que el clúster, elija una de las siguientes soluciones para continuar:

a. Cree un clúster nuevo y elija una subred para el clúster que esté en una zona de disponibilidad en la que el tipo de instancia que ha seleccionado esté disponible y sea compatible con Amazon EMR.

b. Cree un clúster nuevo en la misma región y EC2 subred de Amazon que el clúster en el que se produjo el error, pero con un tipo de instancia que Amazon EMR admita en esa ubicación.

Para ver una lista de tipos de instancia y regiones compatibles con Amazon EMR, consulte Tipos de instancias admitidas con Amazon EMR. Para comparar las capacidades de los tipos de instancias, consulta Tipos de EC2 instancias de Amazon.

Errores de recursos durante las operaciones del clústeres de Amazon EMR

Los siguientes errores suelen ser causados habitualmente por recursos limitados en el clúster. Las instrucciones describen cada error y proporcionan ayuda para la solución de problemas.

Temas

- El clúster de Amazon EMR finaliza con NO_SLAVE_LEFT y los nodos básicos con FAILED_BY_MASTER
- Error de clúster de Amazon EMR: “No se puede replicar el bloque, solo se pudo replicar a cero nodos”.
- Error de clúster de Amazon EMR: EC2 SE HA SUPERADO LA CUOTA
- Error de clúster de Amazon EMR: “Demasiados errores de búsqueda”
- Error de clúster de Amazon EMR: “El archivo solo pudo ser replicado a 0 nodos en lugar de 1”
- Error de clúster de Amazon EMR: “Nodos en la lista de denegación”
- Limitación de errores con un clúster de Amazon EMR
- Error de clúster de Amazon EMR: “No se admite el tipo de instancia”
- Error de clúster de Amazon EMR: EC2 no tiene capacidad
- Error de clúster de Amazon EMR: “Error del factor de replicación de HDFS”
- Error de clúster de Amazon EMR: “Error de espacio insuficiente en HDFS”

El clúster de Amazon EMR finaliza con NO_SLAVE_LEFT y los nodos básicos con FAILED_BY_MASTER

Normalmente, esto ocurre porque la protección de terminación está deshabilitada y todos los nodos secundarios superan la capacidad de almacenamiento en disco especificada por el umbral de utilización máximo en la configuración de clasificación `yarn-site`, que corresponde al archivo `yarn-site.xml`. Este valor es el 90 % de forma predeterminada. Cuando la utilización del disco de un nodo principal supera el umbral de utilización, el servicio de NodeManager salud de YARN informa del nodo como`UNHEALTHY`. Mientras esté en este estado, Amazon EMR lo incluye en la lista de denegados y no le asigna contenedores YARN. Si el nodo permanece en mal estado durante 45 minutos, Amazon EMR marca la EC2 instancia de Amazon asociada para su finalización como. `FAILED_BY_MASTER` Cuando todas las EC2 instancias de Amazon asociadas a los nodos principales están marcadas para su finalización, el clúster termina con el estado `NO_SLAVE_LEFT` porque no hay recursos para ejecutar tareas.

Sobrepasar la utilización del disco en un nodo secundario podría causar una reacción en cadena. Si un único nodo supera el umbral de utilización del disco debido a HDFS, es posible que otros nodos estén también cerca del umbral. El primer nodo supera el umbral de uso del disco, por lo que Amazon EMR lo agrega a la lista de denegados. Esto aumenta la carga de uso del disco en los nodos restantes, ya que estos comienzan a replicar entre ellos los datos HDFS que perdieron en el nodo incluido en la lista de denegados. Uno por uno, los nodos van adoptando el estado `UNHEALTHY` de la misma manera, y el clúster finalmente termina.

Prácticas recomendadas y recomendaciones

Configurar el hardware del clúster con almacenamiento suficiente

Al crear un clúster, asegúrese de que haya suficientes nodos secundarios y de que cada uno tenga un almacén de instancias y volúmenes de almacenamiento de EBS para HDFS apropiados. Para obtener más información, consulte Cálculo de la capacidad de HDFS requerida de un clúster. También puede añadir instancias secundarias a grupos de instancias existentes de forma manual o mediante el escalado automático. Las instancias nuevas tienen la misma configuración de almacenamiento que el resto de las instancias del grupo. Para obtener más información, consulte Utilice el escalado de clústeres de Amazon EMR para adaptarse a las cargas de trabajo cambiantes.

Cómo habilitar la protección contra la terminación

Habilite la protección de terminación. De esta forma, si un nodo principal aparece en la lista de denegados, puedes conectarte a la EC2 instancia de Amazon asociada mediante SSH para

solucionar problemas y recuperar datos. Si habilita la protección de terminación, tenga en cuenta que Amazon EMR no reemplaza la instancia de Amazon EC2 por una nueva instancia. Para obtener más información, consulte Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental.

Cree una alarma para la métrica MRUnhealthy Nodes CloudWatch

Esta métrica indica el número de nodos que tienen el estado `UNHEALTHY`. Es equivalente a la métrica `mapred.resourcemanager.NoOfUnhealthyNodes` de YARN. Puede configurar una notificación para esta alarma que le avise de los nodos en mal estado 45 minutos antes de que se agote el tiempo de espera. Para obtener más información, consulte Supervisión de las métricas de Amazon EMR con CloudWatch.

Retocar la configuración mediante yarn-site

Las opciones mostradas a continuación se pueden ajustar de acuerdo con los requisitos de la aplicación. Por ejemplo, es posible que desee aumentar el umbral de utilización del disco si un nodo adopta el estado `UNHEALTHY` aumentando el valor de `yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-percentage`.

Puede establecer estos valores al crear un clúster mediante la clasificación de configuración `yarn-site`. Para obtener más información, consulte Configuración de aplicaciones en la Guía de publicación de Amazon EMR. También puedes conectarte a las EC2 instancias de Amazon asociadas a los nodos principales mediante SSH y, a continuación, añadir los valores `/etc/hadoop/conf.empty/yarn-site.xml` mediante un editor de texto. Tras realizar el cambio, debes reiniciarlo hadoop-yarn-nodemanager como se muestra a continuación.

> **Important**
>
> Al reiniciar el NodeManager servicio, los contenedores YARN activos se eliminan a menos que `yarn.nodemanager.recovery.enabled` estén configurados para `true` usar la clasificación de `yarn-site` configuración al crear el clúster. Asimismo, debe especificar el directorio en el que se va a almacenar el estado del contenedor mediante la propiedad `yarn.nodemanager.recovery.dir`.

```
sudo /sbin/stop hadoop-yarn-nodemanager
sudo /sbin/start hadoop-yarn-nodemanager
```

Para obtener más información sobre las propiedades `yarn-site` actuales y sus valores predeterminados, consulte página relacionada con la configuración predeterminada de YARN en la documentación de Apache Hadoop.

Propiedad	Valor predeterminado	Descripción
yarn.nodemanager. disk-health-checker.interval-ms	120 000	La frecuencia (en segundos) con la que se ejecuta el comprobador de estado del disco.
yarn.nodemanager. disk-health-checker. min-healthy-disks	0,25	La fracción mínima de la cantidad de discos que deben estar en buen estado NodeManager para lanzar nuevos contenedores. Esto se corresponde con yarn.node manager.local-dirs (de forma predeterminada, `/mnt/yarn` en Amazon EMR) y yarn.node manager.log-dirs (de forma predeterminada, `/var/log/hadoop-yarn/containers` , que está vinculado mediante symlinks a `mnt/var/log/hadoop-yarn/containers` en Amazon EMR).
`yarn.nodemanager.disk-health-checker.max-disk-utilization-per-disk-percentage`	90,0	El porcentaje máximo de utilización del espacio en disco permitido después del cual un disco se marca como dañado. Los valores están comprendidos entre 0,0 y 100,0. Si el valor es mayor o igual a 100, NodeManag

Propiedad	Valor predeterminado	Descripción
		er comprueba si el disco está lleno. Esto se aplica a `yarn-nodemanager.local-dirs` y `yarn.nodemanager.log-dirs` .
`yarn.nodemanager.disk-health-checker.min-free-space-per-disk-mb`	0	El espacio mínimo que debe estar disponible en un disco para que se pueda utilizar. Esto se aplica a `yarn-nodemanager.local-dirs` y `yarn.nodemanager.log-dirs` .

Error de clúster de Amazon EMR: “No se puede replicar el bloque, solo se pudo replicar a cero nodos”.

El error “No se puede replicar el bloque, solo se pudo replicar a cero nodos”. suele ocurrir cuando un clúster no tiene suficiente espacio de almacenamiento de HDFS. Este error se produce cuando se generan más datos en el clúster de los que pueden almacenarse en HDFS. Ve este error solo mientras se está ejecutando el clúster, porque cuando el trabajo termina se libera el espacio de HDFS que se estaba utilizando.

La cantidad de espacio HDFS disponible para un clúster depende del número y el tipo de EC2 instancias de Amazon que se utilicen como nodos principales. Los nodos de tarea no se utilizan para almacenamiento de HDFS. Todo el espacio en disco de cada EC2 instancia de Amazon, incluidos los volúmenes de almacenamiento de EBS adjuntos, está disponible para HDFS. Para obtener más información sobre la cantidad de almacenamiento local para cada tipo de EC2 instancia, consulta Tipos y familias de instancias en la Guía del EC2 usuario de Amazon.

El otro factor que puede influir en la cantidad de espacio de HDFS disponible es el factor de replicación, que es el número de copias de cada bloque de datos que se almacena en HDFS para redundancia. El factor de replicación aumenta con el número de nodos del clúster: existen tres copias de cada bloque de datos para un clúster con 10 o más nodos, 2 copias de cada bloque por un clúster con 4 a 9 nodos y 1 copia (sin redundancia) para clústeres con 3 o menos nodos. El espacio de

HDFS total disponible se divide por el factor de replicación. En algunos casos, como, por ejemplo, aumentando el número de nodos de 9 a 10, el aumento del factor de replicación pueden en realidad hacer que la cantidad de espacio de HDFS disponible se reduzca.

Por ejemplo, un clúster con diez nodos secundarios de tipo m1.large tendría 2 833 GB de espacio disponible para HDFS ((10 nodos X 850 GB por nodo)/factor de replicación de 3).

Si el clúster supera la cantidad de espacio disponible para HDFS, puede añadir más nodos secundarios a su clúster o utilizar la compresión de datos para crear más espacio de HDFS. Si tu clúster se puede detener y reiniciar, puedes considerar usar nodos principales de un tipo de EC2 instancia de Amazon más grande. También puede plantearse la posibilidad de ajustar el factor de replicación. Tenga en cuenta que la disminución del factor de replicación reduce la redundancia de datos de HDFS y la capacidad de su clúster recuperarse frente a bloques de HDFS perdidos o dañados.

Error de clúster de Amazon EMR: EC2 SE HA SUPERADO LA CUOTA

Si recibe un mensaje `EC2 QUOTA EXCEEDED`, se puede deber a varias causas. En función de las diferencias de configuración, los clústeres anteriores pueden tardar entre 5 y 20 minutos en terminar y liberar los recursos asignados. Si aparece un error `EC2 QUOTA EXCEEDED` al intentar lanzar un clúster, puede deberse a que aún no se hayan liberado los recursos de un clúster terminado recientemente. Este mensaje también puede deberse al cambio de tamaño de un grupo o flota de instancias a un tamaño de destino mayor que la cuota de instancias actual para la cuenta. Esto puede ocurrir de forma manual o automática a través de escalado automático.

Tenga en cuenta las opciones siguientes para resolver el problema:

- Siga las instrucciones de AWS Service Quotas en Referencia general de Amazon Web Services para solicitar un aumento del límite de servicio. Para algunos APIs, configurar un CloudWatch evento puede ser una mejor opción que aumentar los límites. Para obtener más información, consulta Cuándo configurar los eventos de EMR en CloudWatch.
- Si uno o varios clústeres en ejecución no se ejecutan según la capacidad, cambie de tamaño los grupos de instancia o reduzca las capacidades de destino en flotas de instancia para clústeres en ejecución.
- Cree clústeres con menos EC2 instancias o con una capacidad de destino reducida.

Error de clúster de Amazon EMR: “Demasiados errores de búsqueda”

La presencia de mensajes de error Too many fetch-failures (Demasiados errores de recuperación) o Error reading task output (Error al leer la salida de la tarea) en los registros de intentos de tareas o de pasos indican que la tarea en ejecución depende de la salida de otra tarea. Esto suele ocurrir cuando una tarea de reducción se pone en cola para ejecutarse y requiere la salida de una o más tareas de asignación y la salida no está disponible aún.

Hay varias razones por las que la salida podría no estar disponible:

- La tarea requisito previo aún se está procesando. Suele ser una tarea de asignación.
- Los datos podrían no estar disponibles debido a una mala conectividad de red si los datos se encuentran en otra instancia.
- Si HDFS se utiliza para recuperar la salida, es posible que exista un problema con HDFS.

La causa más frecuente de este error es que la tarea anterior sigue en procesamiento. Esto es especialmente probable si los errores se producen cuando las tareas de reducción son las primeras que se intentan ejecutar. Puede comprobar si este es el caso revisando el registro syslog para el paso de clúster que devuelve el error. Si el syslog muestra las tareas de asignación y reducción progresando, esto indica que la fase de reducción ha comenzado mientras hay tareas de asignación que no se han completado aún.

Una cosa que hay que buscar en los registros es un porcentaje de progreso de asignación que pasa al 100% y, a continuación, disminuye hasta un valor inferior. Cuando el porcentaje de asignación está al 100%, eso no significa que todas las tareas de asignación se han completado. Simplemente significa que Hadoop está ejecutando todas las tareas de asignación. Si este valor vuelve a disminuir por debajo del 100%, significa que una tarea de asignación ha devuelto un error y, en función de la configuración, Hadoop podría intentar volver a programar la tarea. Si el porcentaje del mapa se mantiene en el 100% de los registros, observe específicamente `RunningMapTasks` las CloudWatch métricas para comprobar si la tarea del mapa aún se está procesando. También puede encontrar esta información a través de la interfaz web de Hadoop en el nodo principal.

Si está viendo este problema, hay varias cosas que puede probar:

- Indique a la fase de reducción que espere más tiempo antes de empezar. Puede hacerlo modificando el ajuste de configuración de Hadoop mapred.reduce.slowstart.completed.maps a un tiempo superior. Para obtener más información, consulte Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR.

- Asigne el recuento de reductores a la capacidad de reductor total del clúster. Esto se hace ajustando la opción de configuración de Hadoop mapred.reduce.tasks para el trabajo.
- Utilice un código de clase de combinador para minimizar el número de salidas que se tienen que recuperar.
- Comprueba que no haya problemas con el EC2 servicio de Amazon que afecten al rendimiento de la red del clúster. Puede hacerlo utilizando el Panel de estado del servicio.
- Revise los recursos de CPU y de memoria de las instancias en su clúster para asegurarse de que su procesamiento de datos no esté desbordando los recursos de los nodos. Para obtener más información, consulte Configuración del hardware y las redes de los clústeres de Amazon EMR.
- Compruebe la versión de la imagen de máquina de Amazon (AMI) utilizada en el clúster de Amazon EMR. Si la versión es de la 2.3.0 a la 2.4.4 incluida, actualice a una versión posterior. Las versiones de AMI en el rango especificado utilizan una versión de Jetty que podría fallar a la hora de entregar la salida desde la fase de asignación. El error de recuperación se produce cuando los reductores no pueden obtener la salida desde la fase de asignación.

  Jetty es un servidor HTTP de código abierto que se utiliza para comunicaciones de equipo a equipo dentro de un clúster de Hadoop.

Error de clúster de Amazon EMR: “El archivo solo pudo ser replicado a 0 nodos en lugar de 1”

Cuando un archivo se escribe en HDFS, se replica a varios nodos secundarios. Si aparece este error, significa que el NameNode daemon no tiene ninguna DataNode instancia disponible en la que escribir datos en HDFS. En otras palabras, la replicación de bloques no se está produciendo. Este error puede deberse a una serie de problemas:

- El sistema de archivos HDFS podría haberse quedado sin espacio. Esta es la causa más probable.
- DataNode es posible que las instancias no estuvieran disponibles cuando se ejecutó el trabajo.
- DataNode Es posible que se haya bloqueado la comunicación de las instancias con el nodo maestro.
- Las instancias del grupo de instancias secundarias podrían no estar disponibles.
- Es posible que falten permisos. Por ejemplo, es posible que el JobTracker daemon no tenga permisos para crear información sobre el registro de tareas.
- La configuración de espacio reservado para una DataNode instancia puede ser insuficiente. Compruebe si este es el caso comprobando la opción de configuración dfs.datanode.du.reserved.

Para comprobar si este problema se debe a que el HDFS se está quedando sin espacio en disco, consulte la `HDFSUtilization` métrica que aparece en CloudWatch. Si este valor es demasiado alto, puede añadir nodos secundarios adicionales en el clúster. Si tiene un clúster y cree que podría quedarse sin espacio en el disco HDFS, puede configurar una alarma CloudWatch para que le avise cuando el valor de supere `HDFSUtilization` un nivel determinado. Para obtener más información, consulte Cambio manual del tamaño de un clúster de Amazon EMR en ejecución y Supervisión de las métricas de Amazon EMR con CloudWatch.

Si el problema no es que el HDFS se quede sin espacio, compruebe los registros, los DataNode NameNode registros y la conectividad de la red para ver si hay otros problemas que pudieran haber impedido que el HDFS replicara los datos. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

Error de clúster de Amazon EMR: “Nodos en la lista de denegación”

El NodeManager daemon es responsable de lanzar y gestionar los contenedores en los nodos principales y de tareas. El NodeManager daemon que se ejecuta en el nodo maestro asigna los contenedores al ResourceManager daemon. ResourceManager Supervisa el NodeManager nodo en un abrir y cerrar de ojos.

Hay un par de situaciones en las que el ResourceManager daemon deniega la lista a NodeManager y la elimina del grupo de nodos disponibles para procesar tareas:

- Si no NodeManager ha enviado ni un latido al ResourceManager daemon en los últimos 10 minutos (600.000 milisegundos). Este periodo de tiempo puede configurarse mediante la opción de configuración `yarn.nm.liveness-monitor.expiry-interval-ms`. Para obtener más información sobre cómo cambiar la configuración de Yarn, consulte Configuración de aplicaciones en la Guía de publicación de Amazon EMR.
- NodeManager comprueba el estado de los discos determinado por `yarn.nodemanager.local-dirs` y. `yarn.nodemanager.log-dirs` Las comprobaciones incluyen permisos y espacio libre en disco (< 90 %). Si un disco no supera la comprobación, NodeManager deja de utilizar ese disco en particular, pero sigue indicando que el estado del nodo es correcto. Si varios discos no pasan la comprobación, el nodo se considera en mal estado ResourceManager y no se asignan nuevos contenedores al nodo.

El maestro de la aplicación también puede denegar la inclusión de un NodeManager nodo en la lista si tiene más de tres tareas fallidas. Puede cambiar esto a un valor superior utilizando el parámetro de configuración `mapreduce.job.maxtaskfailures.per.tracker`. Otras

opciones de configuración que podría cambiar controlan cuántas veces se intenta una tarea antes de marcarla como errónea: `mapreduce.map.max.attempts` para tareas de asignación y `mapreduce.reduce.maxattempts` para tareas de reducción. Para obtener más información sobre cómo cambiar la configuración, consulte Configuración de aplicaciones en la Guía de publicación de Amazon EMR.

Limitación de errores con un clúster de Amazon EMR

Los errores «Limitado desde *Amazon EC2* al lanzar el clúster» y «No se pudieron aprovisionar las instancias debido a la limitación desde» *Amazon EC2* se producen cuando Amazon EMR no puede completar una solicitud porque otro servicio ha limitado la actividad. Amazon EC2 es la fuente más común de errores de regulación, pero otros servicios pueden ser la causa de estos errores. AWS Los límites de servicio se aplican por región para mejorar el rendimiento, y un error de limitación indica que has superado el límite de servicio de tu cuenta en esa región.

Causas posibles

La fuente más común de errores de EC2 limitación de Amazon es el lanzamiento de un gran número de instancias de clúster, por lo que se supera el límite de servicio de EC2 instancias. Las instancias de clúster pueden lanzarse por las siguientes razones:

- Se han creado nuevos clústeres.
- Se ha cambiado manualmente el tamaño de los clústeres. Para obtener más información, consulte Cambio manual del tamaño de un clúster de Amazon EMR en ejecución.
- Los grupos de instancias en un clúster añaden instancias (escalado) como resultado de una regla de escalado automático. Para obtener más información, consulte Descripción de las reglas de escalado automático.
- Las flotas de instancia de un clúster añaden instancias para satisfacer una mayor capacidad de destino. Para obtener más información, consulte Planificación y configuración de flotas de instancias para su clúster de Amazon EMR.

También es posible que la frecuencia o el tipo de solicitud de API que se realiza a Amazon EC2 provoque errores de limitación. Para obtener más información sobre cómo Amazon regula las EC2 solicitudes de API, consulta la tasa de solicitudes de API de consulta en la referencia de EC2 API de Amazon.

Soluciones

Tenga en cuenta las soluciones siguientes:

- Siga las instrucciones de AWS Service Quotas en Referencia general de Amazon Web Services para solicitar un aumento del límite de servicio. Para algunos APIs, configurar un CloudWatch evento puede ser una mejor opción que aumentar los límites. Para obtener más información, consulta Cuándo configurar los eventos de EMR en CloudWatch.
- Si tiene clústeres que se lanzan según una misma programación (por ejemplo, al final de una hora), considere la posibilidad de escalonar las horas de inicio.
- Si tiene clústeres que están dimensionados para picos de demanda y periódicamente tiene problemas de capacidad de la instancia, considere la posibilidad de especificar el escalado automático para añadir y eliminar instancias bajo demanda. De esta forma, las instancias se utilizan de manera más eficiente y, en función del perfil de la demanda, es posible que se soliciten menos instancias en un momento dado en una cuenta. Para obtener más información, consulte Uso del escalado automático con una política personalizada para grupos de instancias en Amazon EMR.

Error de clúster de Amazon EMR: “No se admite el tipo de instancia”

Si creas un clúster y se produce un error con el mensaje de error «El tipo de instancia solicitado no *InstanceType* es compatible con la zona de disponibilidad solicitada», significa que has creado el clúster y especificado un tipo de instancia para uno o más grupos de instancias que no es compatible con Amazon EMR en la región y la zona de disponibilidad en las que se creó el clúster. Amazon EMR puede admitir un tipo de instancia en una zona de disponibilidad de una región y no en otra. La subred que seleccione para un clúster determina la zona de disponibilidad dentro de la región.

Solución

Determine los tipos de instancias disponibles en una zona de disponibilidad mediante el AWS CLI

- Utilice el comando `ec2 run-instances` con la opción `--dry-run`. En el ejemplo siguiente, *m5.xlarge* sustitúyalo por el tipo de instancia que quieres usar, *ami-035be7bafff33b6b6* por la AMI asociada a ese tipo de instancia y *subnet-12ab3c45* por una subred en la zona de disponibilidad que deseas consultar.

```
aws ec2 run-instances --instance-type m5.xlarge --dry-run --image-
id ami-035be7bafff33b6b6 --subnet-id subnet-12ab3c45
```

Para obtener instrucciones sobre cómo encontrar un ID de AMI, consulte Buscar una AMI de Linux. Para encontrar un ID de subred, puede usar el comando describe-subnets.

Para obtener más información sobre cómo descubrir los tipos de instancias disponibles, consulta Buscar un tipo de EC2 instancia de Amazon.

Después de determinar los tipos de instancia disponibles, puede hacer lo siguiente:

- Cree el clúster en la misma región y EC2 subred y elija un tipo de instancia diferente con capacidades similares a las de su elección inicial. Para ver una lista de los tipos de instancia admitidos, consulte Tipos de instancias admitidas con Amazon EMR. Para comparar las capacidades de los tipos de EC2 instancias, consulta Tipos de EC2 instancias de Amazon.
- Elija una subred para el clúster en una zona de disponibilidad en la que el tipo de instancia esté disponible y sea compatible con Amazon EMR.

Mitigue los errores de lanzamiento del clúster de la flota de instancias debido a tipos de instancias principales no compatibles en Amazon EMR

Los nodos principales son esenciales en los clústeres de Amazon EMR. Es posible que se produzca un error de `instance type not supported` en el lanzamiento de un clúster en el que Amazon EMR intente lanzar el clúster en una zona de disponibilidad si el tipo de instancia principal no es compatible. La selección mejorada de zonas de disponibilidad, por ejemplo, los clústeres de flota en Amazon EMR, filtra automáticamente los tipos de instancias principales que especificó en la AZs configuración del clúster y descarta automáticamente los tipos de instancias principales que especificó. Esto significa que Amazon EMR no elegirá una zona de disponibilidad en la que no se admitan los tipos de instancias principales configurados, lo que evita que se produzcan errores en el lanzamiento del clúster debido a tipos de instancias no compatibles.

Para permitir esta mejora, añada el permiso necesario a la política o al rol de servicio de su clúster. La versión más reciente de `AmazonEMRServicePolicy_v2` incluye este permiso, por lo que si utiliza esa política, la mejora ya está disponible. Si utiliza una política o un rol de servicio personalizados, agregue el permiso `ec2:DescribeInstanceTypeOfferings` al lanzar su clúster.

JSON

```
{
```

```
  "Version": "2012-10-17",
  "Statement": [
    {
      "Action": [
        "ec2:DescribeInstanceTypeOfferings"
      ],
      "Effect": "Allow",
      "Resource": [
        "*"
      ],
      "Sid": "AllowEC2Describeinstancetypeofferings"
    }
  ]
}
```

Error de clúster de Amazon EMR: EC2 no tiene capacidad

EC2 Se produce un *InstanceType* error por falta de capacidad al intentar crear un clúster, o añadir instancias a un clúster, en una zona de disponibilidad que ya no tiene el tipo de EC2 instancia especificado. La subred que seleccione para un clúster determina la zona de disponibilidad.

Para crear un clúster, realice alguna de las siguientes operaciones:

- Especificar un tipo de instancia diferente con capacidades similares
- Crear el clúster en una región diferente
- Seleccione una subred en una zona de disponibilidad en la que pueda estar disponible el tipo de instancia que desee.

Para agregar instancias en un clúster en ejecución, haga una de estas acciones:

- Modifique las configuraciones de grupo de instancias o las configuraciones de flota de instancias para agregar tipos de instancia disponibles con capacidades similares. Para ver una lista de los tipos de instancia admitidos, consulte Tipos de instancias admitidas con Amazon EMR. Para comparar las capacidades de los tipos de EC2 instancias, consulta Tipos de EC2 instancias de Amazon.
- Termine el clúster y vuelva a crearlo en una región y zona de disponibilidad en la que el tipo de instancia esté disponible.

Error de clúster de Amazon EMR: “Error del factor de replicación de HDFS”

Al eliminar un nodo básico de un grupo de instancias principal o de una flota de instancias, Amazon EMR podría sufrir un error de replicación de HDFS. Este error se produce cuando se eliminan los nodos básicos y el número de nodos básicos es inferior al factor dfs.replication configurado para el Sistema de archivos distribuido de Hadoop (HDFS). Por lo tanto, Amazon EMR no puede realizar la operación de forma segura. Para determinar el valor predeterminado de la configuración `dfs.replication`, utilice la Configuración HDFS.

Causas posibles

Consulte lo siguiente para conocer las posibles causas del error del factor de replicación de HDFS:

- Si cambia manualmente el tamaño de un grupo de instancias principal o una flota de instancias por debajo del factor configurado `dfs.replication`.
- Sus políticas de escalado administrado o de escalado automático pueden permitir el escalado para reducir la cantidad de nodos básicos por debajo del umbral de `dfs.replication`.
- Este error también puede producirse si Amazon EMR intenta reemplazar un nodo básico en mal estado cuando un clúster tiene el número mínimo de nodos básicos definido por `dfs.replication`.

Soluciones y prácticas recomendadas

Consulte lo siguiente para obtener soluciones y prácticas recomendadas:

- Cuando cambie manualmente el tamaño de un clúster de Amazon EMR, no lo reduzca verticalmente por debajo de `dfs.replication` ya que Amazon EMR no puede completar el cambio de tamaño de forma segura.
- Cuando utilice el escalado administrado o el escalado automático, asegúrese de que la capacidad mínima del clúster no sea inferior al factor `dfs.replication`.
- El número de instancias principales debe ser como mínimo `dfs.replication` más una. Esto garantiza que Amazon EMR pueda reemplazar correctamente un nodo básico en mal estado si habilitó el reemplazo de núcleo en mal estado.

> **Important**
>
> El fallo de un solo nodo básico puede provocar la pérdida de datos del HDFS si usted establece `dfs.replication` en 1. Si su clúster tiene almacenamiento HDFS, le recomendamos que configure el clúster con al menos cuatro nodos principales para las cargas de trabajo de producción a fin de evitar la pérdida de datos y que también establezca el `dfs.replication` factor en al menos 2.

Error de clúster de Amazon EMR: “Error de espacio insuficiente en HDFS”

Si intenta eliminar un nodo básico, puede producirse un error de espacio insuficiente en el Sistema de archivos distribuido de Hadoop (HDFS), pero Amazon EMR no puede completar la operación de forma segura debido a que no queda suficiente espacio en el HDFS. Antes de que Amazon EMR elimine un nodo básico, todos los datos de HDFS del nodo deben transferirse a otros nodos básicos para garantizar la redundancia de los datos. Sin embargo, si no hay suficiente espacio en los otros nodos básicos para la replicación, Amazon EMR no puede realizar una desactivación rápida del nodo.

Causas posibles

Consulte lo siguiente para ver una lista de las posibles causas del error de espacio insuficiente en HDFS:

- Si reduce vertical y manualmente un grupo de instancias básico o una flota de instancias cuando no hay suficiente espacio en HDFS en los nodos restantes para replicar los datos antes de la reducción de escala.
- El escalado administrado o el escalado automático reduce verticalmente un grupo de instancias principal o una flota de instancias cuando no hay suficiente espacio en HDFS para la replicación de datos.
- Amazon EMR intenta reemplazar un nodo principal en mal estado, pero no puede reemplazar el nodo de forma segura debido a la falta de espacio en HDFS.

Soluciones y prácticas recomendadas

Consulte lo siguiente para obtener soluciones y prácticas recomendadas:

- Escale verticalmente la cantidad de nodos básicos de su clúster de Amazon EMR. Si utiliza el escalado administrado o el escalado automático, aumente la capacidad mínima de sus nodos básicos.
- Utilice volúmenes de EBS más grandes para sus nodos básicos al crear su clúster de EMR.
- Elimine los datos HDFS innecesarios del clúster de EMR. Le recomendamos que configure CloudWatch alarmas para monitorear la `HDFSUtilization` métrica de su clúster para saber si su clúster de EMR tiene poco espacio.

Errores de entrada y salida de clúster durante la operaciones de Amazon EMR

Los errores siguientes son habituales en las operaciones de entrada y salida de clúster. Utilice las instrucciones de este tema para solucionar problemas y comprobar su configuración.

Temas

- ¿La ruta hacia Amazon Simple Storage Service (Amazon S3) tiene por lo menos tres barras?
- ¿Está intentando atravesar de forma recursiva directorios de entrada?
- ¿Ya existe el directorio de salida?
- ¿Está intentando especificar un recurso mediante una URL HTTP?
- ¿Está haciendo referencia a un bucket de Amazon S3; con un formato de nombre no válido?
- ¿Tiene problemas para cargar datos hacia o desde Amazon S3?

¿La ruta hacia Amazon Simple Storage Service (Amazon S3) tiene por lo menos tres barras?

Cuando especifique un bucket de Amazon S3, deberá incluir una barra de terminación al final de la URL. Por ejemplo, en lugar de hacer referencia a un bucket como “s3n://amzn-s3-demo-bucket1”, debería utilizar “s3n://amzn-s3-demo-bucket1/”; de lo contrario, Hadoop genera un error en el clúster en la mayoría de los casos.

¿Está intentando atravesar de forma recursiva directorios de entrada?

Hadoop no busca directorios de entrada de forma recursiva para archivos. Si tiene una estructura de directorios como /corpus/01/01.txt, /corpus/01/02.txt, /corpus/02/01.txt, etc. y especifica /corpus/ como parámetro de entrada para el clúster, Hadoop no encuentra los archivos de entrada, ya que el

directorio /corpus/ está vacío y Hadoop no comprueba el contenido de los subdirectorios. Del mismo modo, Hadoop no comprueba recursivamente los subdirectorios de buckets de Amazon S3.

Los archivos de entrada deben estar directamente en el directorio de entrada o bucket de Amazon S3 que especifique, no en subdirectorios.

¿Ya existe el directorio de salida?

Si especifica una ruta de salida que ya existe, Hadoop generará un error en el clúster en la mayoría de los casos. Esto significa que si ejecuta un clúster una vez y, a continuación, lo vuelve a ejecutar con los mismos parámetros, probablemente funcionará la primera vez y, a continuación, nunca más; después de la primera ejecución, existe la ruta de salida y, por lo tanto, hace que todas las ejecuciones sucesivas generen un error.

¿Está intentando especificar un recurso mediante una URL HTTP?

Hadoop no acepta las ubicaciones de los recursos especificados mediante el prefijo http://. No puede hacer referencia a un recurso con una dirección URL HTTP. Por ejemplo, transferir http://mysite/myjar.jar como parámetro JAR provoca que el clúster devuelva un error.

¿Está haciendo referencia a un bucket de Amazon S3; con un formato de nombre no válido?

Si intenta utilizar un nombre de bucket como “amzn-s3-demo-bucket1.1” con Amazon EMR, el clúster devolverá un error porque Amazon EMR requiere que los nombres de bucket sean nombres de host RFC 2396 válidos; el nombre no puede terminar por un número. Además, debido a los requisitos de Hadoop, los nombres de bucket de Amazon S3 que se utilizan con Amazon EMR solo puede contener letras en minúsculas, números, puntos (.) y guiones (-). Para obtener más información acerca de las convenciones de nomenclatura de buckets de Amazon S3, consulte Restricciones y limitaciones de los buckets en la Guía del usuario de Amazon Simple Storage Service.

¿Tiene problemas para cargar datos hacia o desde Amazon S3?

Amazon S3 es el origen de entrada y salida más popular para Amazon EMR. Un error común consiste en tratar Amazon S3 como si fuera un sistema de archivos habitual. Existen diferencias entre Amazon S3 y un sistema de archivos que debe tener en cuenta a la hora de ejecutar el clúster.

- Si se produce un error interno en Amazon S3, la aplicación tiene que gestionarlo correctamente y volver a intentar la operación.

- Si las llamadas a Amazon S3 tardan demasiado tiempo en devolverse, es posible que la aplicación tenga que reducir la frecuencia con la que llama a Amazon S3.
- Enumerar todos los objetos en un bucket de Amazon S3 es una llamada costosa. La aplicación debe minimizar el número de veces que lo hace.

Existen varias formas de mejorar la forma en la que el clúster interactúa con Amazon S3.

- Inicie el clúster con la versión de lanzamiento más reciente de Amazon EMR.
- Utilice S3 DistCp para mover objetos dentro y fuera de Amazon S3. S3 DistCp implementa la gestión de errores, los reintentos y los retrasos para cumplir con los requisitos de Amazon S3. Para obtener más información, consulte Copia distribuida mediante S3. DistCp
- Diseñe la aplicación teniendo en cuenta la consistencia final. Utilice HDFS para el almacenamiento de datos intermedios mientras que el clúster se está ejecutando y Amazon S3 únicamente para entrada de los datos iniciales y salida de los resultados finales.
- Si los clústeres confirmarán 200 o más transacciones por segundo en Amazon S3, contacte con la asistencia técnica para preparar el bucket para que haga más transacciones por segundo y plantéese el uso de las estrategias de partición de claves que se describen en Consejos y trucos de rendimiento de Amazon S3.
- Defina el ajuste de configuración de Hadoop io.file.buffer.size en 65 536. Esto hace que Hadoop dedique menos tiempo a buscar a través de objetos de Amazon S3.
- Plantéese deshabilitar la característica de ejecución especulativa de Hadoop si su clúster experimenta problemas de simultaneidad de Amazon S3. Esto también resulta útil cuando se solucionan problemas de un clúster lento. Para ello, establezca las propiedades `mapreduce.reduce.speculative` `mapreduce.map.speculative` y en `false`. Cuando lance un clúster, podrá establecer estos valores mediante la clasificación de configuración `mapred-env`. Para obtener más información, consulte Configuración de aplicaciones en la Guía de publicación de Amazon EMR.
- Si ejecuta un clúster de Hive, consulte ¿Tiene problemas para cargar datos hacia o desde Amazon S3 en Hive?.

Para obtener información adicional, consulte Prácticas recomendadas de errores de Amazon S3 en la Guía del usuario de Amazon Simple Storage Service.

Errores de permisos durante las operaciones del clúster de Amazon EMR

Los siguientes errores son comunes cuando se utilizan permisos o credenciales.

Temas

- ¿Está transfiriendo las credenciales correctas en SSH?
- Si utilizas IAM, ¿has establecido las EC2 políticas de Amazon adecuadas?

¿Está transfiriendo las credenciales correctas en SSH?

Si no puede utilizar SSH para conectarse al nodo principal, se trata muy probablemente de un problema con sus credenciales de seguridad.

En primer lugar, compruebe que el archivo .pem que contiene su clave SSH disponga de los permisos adecuados. Puede utilizar chmod para cambiar los permisos de su archivo .pem tal y como se muestra en el siguiente ejemplo, donde debería sustituir mykey.pem por el nombre de su propio archivo .pem.

```
chmod og-rwx mykey.pem
```

La segunda posibilidad es que no se esté utilizando el par de claves que especificó al crear el clúster. Esto es fácil de hacer si ha creado varios pares de claves. Consulte los detalles del clúster en la consola de Amazon EMR (o utilice la opción `--describe` de la CLI) para el nombre del par de claves que se especificó cuando se creó el clúster.

Una vez que haya verificado que está utilizando el par de claves correcto y que los permisos se han definido correctamente en el archivo .pem, puede utilizar el siguiente comando para utilizar SSH para conectarse al nodo maestro, donde debería sustituir mykey.pem por el nombre de su archivo .pem y `hadoop@ec2-01-001-001-1.compute-1.amazonaws.com` por el nombre de DNS público del nodo principal (disponible a través de la opción `--describe` en la CLI o a través de la consola de Amazon EMR).

> **Important**
>
> Debe utilizar el nombre de inicio de sesión `hadoop` cuando se conecte a un nodo de clúster de Amazon EMR; de lo contrario, es posible que se produzca un error similar a `Server refused our key`.

```
ssh -i mykey.pem hadoop@ec2-01-001-001-1.compute-1.amazonaws.com
```

Para obtener más información, consulte Conexión al nodo principal del clúster de Amazon EMR mediante SSH.

Si utilizas IAM, ¿has establecido las EC2 políticas de Amazon adecuadas?

Dado que Amazon EMR utiliza EC2 las instancias como nodos, los usuarios de Amazon EMR también deben tener determinadas EC2 políticas de Amazon establecidas para que Amazon EMR pueda gestionar esas instancias en nombre del usuario. Si no ha establecido los permisos necesarios, Amazon EMR devuelve el error: «la cuenta no está autorizada a llamar EC2».

Para obtener más información sobre las EC2 políticas de Amazon que su cuenta de IAM debe configurar para ejecutar Amazon EMR, consulte. Cómo funciona Amazon EMR con IAM

Errores de clúster de Hive

Normalmente, puede encontrar la causa de un error de Hive en el archivo `syslog`, para el que tiene un enlace en el panel Steps (Pasos). Si no puede determinar el problema allí, consulte el mensaje de error de intento de tareas de Hadoop. Encontrará un enlace al mismo en el panel Task Attempts (Intentos de tareas).

Los siguientes errores son comunes en los clústeres de Hive.

Temas

- ¿Está utilizando la última versión de Hive?
- ¿Ha detectado un error de sintaxis en el script de Hive?
- ¿Ha devuelto error un trabajo al ejecutarlo de forma interactiva?

- ¿Tiene problemas para cargar datos hacia o desde Amazon S3 en Hive?

¿Está utilizando la última versión de Hive?

La última versión de Hive presenta todas las revisiones actuales y correcciones de errores y podría resolver el problema.

¿Ha detectado un error de sintaxis en el script de Hive?

Si un paso devuelve un error, examine el archivo `stdout` de los registros para el paso que se ejecutó en el script de Hive. Si el error no se encuentra allí, examine el archivo `syslog` de los registros del intento de tarea que ha devuelto error. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

¿Ha devuelto error un trabajo al ejecutarlo de forma interactiva?

Si ejecuta Hive de forma interactiva en el nodo principal y el clúster ha fallado, vea las entradas `syslog` en el registro de intento de tarea para el intento de tarea fallido. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

¿Tiene problemas para cargar datos hacia o desde Amazon S3 en Hive?

Si tiene problemas para tener acceso a los datos en Amazon S3, compruebe antes las causas posibles incluidas en ¿Tiene problemas para cargar datos hacia o desde Amazon S3?. Si ninguno de estos problemas es la causa, tenga en cuenta las siguientes opciones específicas de Hive.

- Asegúrese de utilizar la última versión de Hive que presenta todas las revisiones actuales y correcciones de errores que podría resolver el problema. Para obtener más información, consulte Apache Hive.
- El uso de `INSERT OVERWRITE` requiere mostrar el contenido del bucket o carpeta de Amazon S3. Se trata de una operación costosa. Si es posible, elimine manualmente la ruta en lugar de que Hive enumere y elimine los objetos existentes.
- Si utiliza versiones de Amazon EMR anteriores a la 5.0, puede utilizar el comando siguiente de HiveQL para guardar previamente en caché los resultados de una operación de listado de Amazon S3 localmente en el clúster:

```
set hive.optimize.s3.query=true;
```

- Utilice las particiones estáticas donde sea posible.
- En algunas versiones de Hive y Amazon EMR, es posible que el uso de ALTER TABLES devuelva un error debido a que la tabla se almacena en una ubicación distinta a la esperada por Hive. La solución consiste en añadir o actualizar lo siguiente en `/home/hadoop/conf/core-site.xml`:

```
<property>
    <name>fs.s3n.endpoint</name>
    <value>s3.amazonaws.com</value>
</property>
```

Errores de la VPC durante las operaciones del clúster de Amazon EMR

Los siguientes errores son comunes a la configuración de VPC en Amazon EMR.

Temas

- Configuración de subredes no válida
- Falta el conjunto de opciones de DHCP
- Errores de permisos
- Errores que dan lugar a START_FAILED
- Se agrupa y no se puede iniciar Terminated with errors NameNode

Configuración de subredes no válida

En la página Cluster Details (Detalles del clúster), en el campo Status (Estado), ve un error similar al siguiente:

`The subnet configuration was invalid: Cannot find route to InternetGateway in main RouteTable` *`rtb-id`* `for vpc` *`vpc-id`*.

Para solucionar este problema, debe crear una gateway de Internet y asociarla a la VPC. Para obtener más información, consulte Adding an internet gateway to your VPC (Cómo añadir una gateway de Internet a la VPC).

De forma alternativa, compruebe que ha configurado la VPC con las opciones Enable DNS resolution (Habilitar resolución de DNS) y Enable DNS hostname support (Habilitar el soporte de nombres de host DNS) habilitadas. Para obtener más información, consulte Utilización de DNS con su VPC.

Falta el conjunto de opciones de DHCP

Puede ver un error de paso en el registro del sistema (syslog) del clúster con un error similar al siguiente:

```
ERROR org.apache.hadoop.security.UserGroupInformation
(main): PriviledgedActionException as:hadoop (auth:SIMPLE)
cause:java.io.IOException:
org.apache.hadoop.yarn.exceptions.ApplicationNotFoundException: Application
with id 'application_id' doesn't exist in RM.
```

o

```
ERROR org.apache.hadoop.streaming.StreamJob (main): Error Launching job :
org.apache.hadoop.yarn.exceptions.ApplicationNotFoundException: Application
with id 'application_id' doesn't exist in RM.
```

Para solucionar este problema, debe configurar una VPC que incluya un conjunto de opciones de DHCP cuyos parámetros se hayan definido en los siguientes valores:

> **Note**
>
> Si usa la región AWS GovCloud (EE. UU. Oeste), establezca el nombre de dominio **`us-gov-west-1.compute.internal`** en lugar del valor utilizado en el siguiente ejemplo.

- domain-name = **`ec2.internal`**

  Use **`ec2.internal`** si su región es Este de EE. UU. (Norte de Virginia). Para otras regiones, utilice. *`region-name`* **`.compute.internal`** Por ejemplo en us-west-2, utilice domain-name=**`us-west-2.compute.internal`**.
- domain-name-servers = **`AmazonProvidedDNS`**

Para obtener más información, consulte Conjuntos de opciones de DHCP.

Errores de permisos

Un error en el registro `stderr` de un paso indica que un recurso de Amazon S3 no tiene los permisos adecuados. Se trata de un error 403 y su aspecto es:

```
Exception in thread "main" com.amazonaws.services.s3.model.AmazonS3Exception: Access
 Denied (Service: Amazon S3; Status Code: 403; Error Code: AccessDenied; Request
 ID: REQUEST_ID
```

Si ActionOnFailure se establece en`TERMINATE_JOB_FLOW`, esto provocará que el clúster termine con el estado,`SHUTDOWN_COMPLETED_WITH_ERRORS`.

Algunas formas de solucionar este problema son:

- Si está utilizando una política de bucket de Amazon S3 dentro de una VPC, asegúrese de proporcionar acceso a todos los buckets. Para ello, cree un punto de conexión de VPC y seleccione Permitir todo en la opción Política al crear el punto de conexión.
- Asegúrese de que las políticas asociadas con recursos de S3 incluyan la VPC en la que lanzar el clúster.
- Pruebe a ejecutar el siguiente comando desde el clúster para verificar que puede acceder al bucket

  ```
  hadoop fs -copyToLocal s3://path-to-bucket /tmp/
  ```

- Puede obtener información más específica sobre la depuración definiendo el parámetro `log4j.logger.org.apache.http.wire` en `DEBUG` en el archivo `/home/hadoop/conf/log4j.properties` en el clúster. Puede comprobar el archivo de registro `stderr` después de intentar acceder al bucket desde el clúster. El archivo de registro proporcionará información más detallada:

  ```
  Access denied for getting the prefix for bucket - us-west-2.elasticmapreduce with
   path samples/wordcount/input/
  15/03/25 23:46:20 DEBUG http.wire: >> "GET /?prefix=samples%2Fwordcount%2Finput
  %2F&delimiter=%2F&max-keys=1 HTTP/1.1[\r][\n]"
  15/03/25 23:46:20 DEBUG http.wire: >> "Host: us-
  west-2.elasticmapreduce.s3.amazonaws.com[\r][\n]"
  ```

Errores que dan lugar a **`START_FAILED`**

Antes de la AMI 3.7.0, para los VPCs casos en los que se especificaba un nombre de host, Amazon EMR mapeaba los nombres de host internos de la subred con direcciones de dominio personalizadas de la siguiente manera:. `ip-`*`X.X.X.X.customdomain.com`*`.tld` Por ejemplo, si el nombre de host fuera `ip-10.0.0.10` y la VPC tuviera la opción de nombre de dominio

definida en customdomain.com, el nombre de host resultante asignado por Amazon EMR sería `ip-10.0.1.0.customdomain.com`. Se añade una entrada en `/etc/hosts` para resolver el nombre de host a 10.0.0.10. Este comportamiento se cambia con la AMI 3.7.0 y ahora Amazon EMR respeta completamente la configuración de DHCP de la VPC. Anteriormente, los clientes también podrían utilizar una acción de arranque para especificar un mapeo de nombre de host.

Si desea conservar este comportamiento, debe proporcionar la DNS y reenviar la configuración de resolución que necesita para el dominio personalizado.

Se agrupa y no se puede iniciar **`Terminated with errors`** NameNode

Al lanzar un clúster de EMR en una VPC que hace uso de un nombre de dominio de DNS personalizado, el clúster podría devolver el siguiente mensaje de error en la consola:

```
Terminated with errors  On the master instance(instance-id), bootstrap action 1
 returned a  non-zero return code
```

El error se debe a que NameNode no se pudo iniciar. Esto provocará el siguiente error en los NameNode registros, cuyo URI de Amazon S3 tiene el siguiente formato`s3://amzn-s3-demo-bucket/logs/cluster-id/daemons/master instance-id/hadoop-hadoop-namenode-master node hostname.log.gz`:

```
2015-07-23 20:17:06,266 WARN
      org.apache.hadoop.hdfs.server.namenode.FSNamesystem (main): Encountered
 exception
      loading fsimage  java.io.IOException: NameNode is not formatted.
      at

 org.apache.hadoop.hdfs.server.namenode.FSImage.recoverTransitionRead(FSImage.java:212)
           at

 org.apache.hadoop.hdfs.server.namenode.FSNamesystem.loadFSImage(FSNamesystem.java:1020)
           at

 org.apache.hadoop.hdfs.server.namenode.FSNamesystem.loadFromDisk(FSNamesystem.java:739)
           at
      org.apache.hadoop.hdfs.server.namenode.NameNode.loadNamesystem(NameNode.java:537)
           at
      org.apache.hadoop.hdfs.server.namenode.NameNode.initialize(NameNode.java:596)

      at  org.apache.hadoop.hdfs.server.namenode.NameNode.<init>(NameNode.java:765)
           at
```

```
      org.apache.hadoop.hdfs.server.namenode.NameNode.<init>(NameNode.java:749)
      at

 org.apache.hadoop.hdfs.server.namenode.NameNode.createNameNode(NameNode.java:1441)
           at
      org.apache.hadoop.hdfs.server.namenode.NameNode.main(NameNode.java:1507)
```

Esto se debe a un posible problema por el que una EC2 instancia puede tener varios conjuntos de nombres de dominio totalmente cualificados al lanzar clústeres de EMR en una VPC, lo que utiliza tanto un servidor DNS proporcionado como un servidor DNS AWS personalizado proporcionado por el usuario. Si el servidor de DNS proporcionado por el usuario no ofrece ningún registro de puntero (PTR) para ningún registro A utilizado para designar nodos en un clúster de EMR, los clústeres devolverán un error al iniciarse cuando se configuran de esta manera. La solución consiste en añadir 1 registro PTR por cada registro A que se cree cuando se lanza una EC2 instancia en cualquiera de las subredes de la VPC.

Errores del clúster de Amazon EMR de streaming

Normalmente, puede encontrar la causa de un error de streaming en un archivo `syslog`. Encontrará un enlace al mismo en el panel Steps (Pasos).

Los siguientes errores son comunes a los clústeres de streaming.

Temas

- ¿Los datos que se envían al mapeador están en formato equivocado?
- ¿Se agota el tiempo de espera del script?
- ¿Está transfiriendo un argumento de streaming no válido?
- ¿El script se cierra con un error?

¿Los datos que se envían al mapeador están en formato equivocado?

Para comprobar si este es el caso, busque un mensaje de error en el archivo `syslog` un intento de tarea con error en los registros de intento de tareas. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

¿Se agota el tiempo de espera del script?

El tiempo de espera predeterminado para un script de mapeador o reductor es 600 segundos. Si el script tarda más tiempo, el intento de tarea devolverá un error. Puede comprobar si es así

comprobando el archivo `syslog` de un intento de tarea con error en los registros de intento de tareas. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

Puede cambiar el límite de tiempo estableciendo un nuevo valor para el ajuste de configuración de `mapred.task.timeout`. Esta configuración especifica el número de milisegundos tras el que Amazon EMR terminará una tarea que no tiene entrada de lectura, salida de escritura o ha actualizado su cadena de estado. Puede actualizar este valor transfiriendo un argumento de streaming adicional `-jobconf mapred.task.timeout=800000`.

¿Está transfiriendo un argumento de streaming no válido?

Hadoop Streaming admite únicamente los siguientes argumentos. Si transfiere argumentos distintos de los que se indican a continuación, el clúster devolverá un error.

```
-blockAutoGenerateCacheFiles
-cacheArchive
-cacheFile
-cmdenv
-combiner
-debug
-input
-inputformat
-inputreader
-jobconf
-mapper
-numReduceTasks
-output
-outputformat
-partitioner
-reducer
-verbose
```

Además, Hadoop Streaming solo reconoce argumentos transferidos mediante sintaxis de Java; es decir, precedidos de un único guion. Si transfiere argumentos precedidos de un guion doble, el clúster fallará.

¿El script se cierra con un error?

Si su script de mapeador o reductor termina con un error, puede localizar el error en el archivo `stderr` de los registros de intento de tarea del intento de tarea que ha devuelto error. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

Amazon EMR: errores de clúster JAR personalizados

Los siguientes errores son comunes en los clústeres JAR personalizados.

Temas

- ¿Su JAR lanza una excepción antes de crear un trabajo?
- ¿Su JAR lanza un error dentro de una tarea de asignación?

¿Su JAR lanza una excepción antes de crear un trabajo?

Si el programa principal de su JAR personalizado arroja una excepción al crear el trabajo de Hadoop, el mejor lugar consiste donde buscar es el archivo `syslog` de los registros de pasos. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

¿Su JAR lanza un error dentro de una tarea de asignación?

Si su JAR personalizado y mapeador lanzan una excepción al procesar los datos de entrada, el mejor lugar donde buscar es el archivo `syslog` de los registros de intento de tarea. Para obtener más información, consulte Visualización de los archivos de registro de Amazon EMR.

Errores de Amazon EMR AWS GovCloud (EE. UU. - Oeste)

La región AWS GovCloud (EE. UU. oeste) se diferencia de otras regiones en cuanto a la seguridad, la configuración y los ajustes predeterminados. Como resultado, utilice la siguiente lista de comprobación para solucionar los errores de Amazon EMR específicos de la región AWS GovCloud (EE. UU. Oeste) antes de utilizar recomendaciones de solución de problemas más generales.

- Compruebe que sus roles de IAM estén correctamente configurados. Para obtener más información, consulte Configuración de los roles de servicio de IAM de los permisos de Amazon EMR para los servicios y recursos de AWS.
- Asegúrese de que la configuración de la VPC haya configurado correctamente los parámetros de resolution/hostname compatibilidad con DNS, Internet Gateway y conjunto de opciones de DHCP.

Para obtener más información, consulte Errores de la VPC durante las operaciones del clúster de Amazon EMR.

Si estos pasos no resuelven el problema, continúe con los pasos para la resolución de los errores de Amazon EMR comunes. Para obtener más información, consulte Recopilaciones de errores comunes en Amazon EMR.

Buscar un clúster que falta

Si el clúster no aparece en la lista de consolas o en la API `ListClusters`, compruebe lo siguiente:

- Confirme que la antigüedad del clúster desde el momento de su finalización es inferior a dos meses. Amazon EMR conserva la información de metadatos sobre clústeres completados, gratuitamente, durante dos meses. No puede eliminar los clústeres completados de la consola; en su lugar, Amazon EMR purga los clústeres completados automáticamente después de dos meses.
- Confirme que tiene los permisos de rol para ver el clúster.
- Confirme que está viendo el mismo Región de AWS lugar donde reside el clúster.

Solución de problemas de un clúster de Amazon EMR que ha fallado con un código de error

Esta sección le muestra el proceso de resolución de problemas de un clúster que ha generado un error. Esto significa que el clúster terminó con un código de error.

> **Note**
>
> Cuando un clúster de EMR termina con un error, `ListClusters` APIs devuelve un código de error `DescribeCluster` y un mensaje de error. En el caso de algunos errores del clúster, la matriz de datos `ErrorDetail` también puede ayudarle a solucionar el error. Para obtener más información, consulte Códigos de error con ErrorDetail información en Amazon EMR.

Si el clúster se ejecuta, pero tarda mucho en devolver resultados, consulte Solución de problemas de un clúster de Amazon EMR lento.

Temas

- Paso 1: recopile datos sobre el problema con el clúster de Amazon EMR
- Paso 2: comprobar el entorno
- Paso 3: comprobar el último cambio de estado
- Paso 4: examine los archivos de registro de Amazon EMR
- Paso 5: compruebe el clúster de Amazon EMR de forma pormenorizada

Paso 1: recopile datos sobre el problema con el clúster de Amazon EMR

El primer paso para solucionar problemas de un clúster es recopilar información sobre lo que ha fallado y el estado y la configuración actuales del clúster. Esta información se utilizará en los siguientes pasos para confirmar o descartar las posibles causas del problema.

Definir el problema

Una definición clara del problema es el primer punto de partida. Algunas preguntas que debe hacerse:

- ¿Qué esperaba que sucediera? ¿Qué pasó en su lugar?
- ¿Cuándo se produjo este problema por primera vez? ¿Con qué frecuencia ha ocurrido desde entonces?
- ¿Ha cambiado algo en la forma en que configuro o ejecuto mi clúster?

Detalles de clúster

Los siguientes detalles del clúster son útiles para ayudar a detectar problemas. Para obtener más información sobre cómo recopilar esta información, consulte Visualización del estado y los detalles del clúster de Amazon EMR.

- El identificador del clúster. (También se denomina identificador de flujo de tareas).
- Región de AWS y la zona de disponibilidad en la que se lanzó el clúster.
- El estado del clúster, incluidos los detalles del último cambio de estado.
- Tipo y número de EC2 instancias especificados para los nodos principal, principal y de tareas.

Paso 2: comprobar el entorno

Amazon EMR opera como parte de un ecosistema de servicios web y software de código abierto. Lo que afecta a dichas dependencias pueden influir en el rendimiento de Amazon EMR.

Temas

- Comprobar las interrupciones de servicio
- Comprobar los límites de uso
- Comprobar la versión
- Comprobar la configuración de subredes de Amazon VPC

Comprobar las interrupciones de servicio

Amazon EMR utiliza varios servicios de Amazon Web Services internamente. Ejecuta servidores virtuales en Amazon EC2, almacena datos y scripts en Amazon S3 e informa de las métricas a CloudWatch. Los eventos que interrumpen estos servicios son poco frecuentes, pero, cuando se producen, pueden provocar problemas en Amazon EMR.

Antes de continuar, compruebe el Panel de estado del servicio. Compruebe la región en la que lanzó el clúster para ver si hay interrupciones en alguno de estos servicios.

Comprobar los límites de uso

Si va a lanzar un clúster grande, ha lanzado varios clústeres simultáneamente o es un usuario que comparte uno Cuenta de AWS con otros usuarios, es posible que el clúster haya fallado porque ha superado el límite de AWS servicio.

Amazon EC2 limita el número de instancias de servidores virtuales que se ejecutan en una sola AWS región a 20 instancias reservadas o bajo demanda. Si lanza un clúster con más de 20 nodos o lanza un clúster que hace que el número total de EC2 instancias activas en la Cuenta de AWS suya supere las 20, el clúster no podrá lanzar todas las EC2 instancias que necesita y podría fallar. Cuando esto ocurre, Amazon EMR devuelve un error `EC2 QUOTA EXCEEDED`. Para solicitar que se AWS aumente el número de EC2 instancias que puede ejecutar en su cuenta, envíe una solicitud de aumento del límite de EC2 instancias de Amazon.

Otro factor que puede provocar que supere los límites de uso es el retraso que transcurre entre la finalización de un clúster y el momento en que libera todos sus recursos. En función de las diferencias de configuración, un clúster puede tardar entre 5 y 20 minutos terminar por completo

y liberar los recursos asignados. Si aparece un error `EC2 QUOTA EXCEEDED` al intentar lanzar un clúster, puede deberse a que aún no se hayan liberado los recursos de un clúster terminado recientemente. En este caso, puedes solicitar que se aumente tu EC2 cuota de Amazon o esperar veinte minutos y volver a lanzar el clúster.

Amazon S3 limita el número de buckets creados en una cuenta a 100. Si el clúster crea un bucket nuevo que supera este límite, se producirá un error en la creación del bucket y es posible que el clúster falle.

Comprobar la versión

Compare la etiqueta de versión que ha usado para lanzar el clúster con la última versión de Amazon EMR. Cada versión de Amazon EMR incorpora mejoras, como nuevas aplicaciones, características, parches y errores corregidos. Puede que el problema que afecta a su clúster ya se haya solucionado en la última versión. Si es posible, vuelva a ejecutar el clúster con la versión más reciente.

Comprobar la configuración de subredes de Amazon VPC

Si el clúster se lanzó en una subred de Amazon VPC, la subred debe configurarse como se describe en Configuración de redes en una VPC para Amazon EMR. Además, compruebe que la subred en la que lanza el clúster tenga suficientes direcciones IP elásticas libres para asignar una a cada nodo del clúster.

Paso 3: comprobar el último cambio de estado

El último cambio de estado ofrece información sobre qué ocurrió la última vez que el clúster cambió de estado. Esto a menudo contiene información capaz de indicar lo que ha funcionado mal cuando el clúster cambia su estado a `FAILED`. Por ejemplo, si lanza un clúster de streaming y especifica una ubicación de salida que ya existe en Amazon S3, se producirá el error con un último cambio de estado “El directorio de salida de streaming ya existe” en el clúster.

Puede localizar el valor del último cambio de estado desde la consola consultando el panel de detalles del clúster, desde la CLI utilizando los argumentos `list-steps` o `describe-cluster` o desde el API utilizando las acciones `DescribeCluster` y `ListSteps`. Para obtener más información, consulte Visualización del estado y los detalles del clúster de Amazon EMR.

Paso 4: examine los archivos de registro de Amazon EMR

El siguiente paso consiste en examinar los archivos de registro para encontrar un código de error u otro indicio del problema que ha sufrido el clúster. Para obtener información sobre los archivos de

registro disponibles, dónde encontrarlos y cómo verlos, consulte Visualización de los archivos de registro de Amazon EMR.

Es posible que sea necesario un poco de trabajo de investigación para determinar qué pasó. Hadoop ejecuta los trabajos en los intentos de tareas en varios nodos del clúster. Amazon EMR puede iniciar intentos de tareas especulativos y terminar los demás intentos de tareas que no se completen primero. Esto genera una actividad importante que se registra en los archivos de registro del controlador, stderr y syslog a medida que se produce. Además, se ejecutan varios intentos de tareas simultáneamente, pero un archivo de registro solo puede mostrar los resultados de forma lineal.

Para comenzar, compruebe los registros de acciones de arranque para ver si hay errores o cambios de configuración inesperados durante el lanzamiento del clúster. A partir de ahí, consulte los registros de pasos para identificar los trabajos de Hadoop lanzados como parte de un paso con errores. Examine los registros de trabajos de Hadoop para identificar los intentos fallidos de tareas. El registro de intentos de tarea contendrá detalles sobre la causa del error de un intento de tarea.

En las siguientes secciones, se describe cómo utilizar los distintos archivos de registro para identificar errores en el clúster.

Comprobar los registros de acción de arranque

Las acciones de arranque ejecutan scripts en el clúster a medida que se lanza. Por lo general, se utilizan para instalar software adicional en el clúster o para modificar los valores predeterminados de los valores de configuración. La comprobación de estos registros puede proporcionar información sobre los errores que se produjeron durante la configuración del clúster, así como sobre los cambios en los ajustes de configuración que podrían afectar al rendimiento.

Comprobar los registros de pasos

Hay cuatro tipos de registros de pasos.

- controlador: contiene archivos generados por Amazon EMR (Amazon EMR) que se deben a errores encontrados al intentar ejecutar el paso. Si se produce un error en el paso durante la carga, puede encontrar el registro de seguimiento de la pila en este registro. Aquí se describen con frecuencia los errores al cargar la aplicación o al acceder a ella, así como los errores que faltan en el archivo de asignación.
- stderr: contiene los mensajes de error que se produjeron al procesar el paso. Los errores de carga de la aplicación se describen a menudo aquí. En ocasiones, este registro contiene un seguimiento de pila.

- stdout: contiene el estado generado por los ejecutables de asignación y reducción. Los errores de carga de la aplicación se describen a menudo aquí. En ocasiones, este registro contiene mensajes de error de la aplicación.
- syslog: contiene registros de software ajeno a Amazon, como Apache y Hadoop. Los errores de streaming suelen describirse aquí.

Compruebe si hay errores obvios en stderr. Si stderr muestra una lista corta de errores, el paso se detuvo rápidamente y se produjo un error. En la mayoría de los casos, esto se debe a un error en las aplicaciones de asignación y reducción que se ejecutan en el clúster.

Examine las últimas líneas del controlador y de syslog en busca de avisos de errores. Siga cualquier aviso sobre tareas con errores, especialmente si dice “Trabajo con errores”.

Comprobar los registros de intento de tarea

Si el análisis anterior de los registros de pasos reveló una o más tareas fallidas, investigue los registros de los intentos de tareas correspondientes para obtener información de error más detallada.

Paso 5: compruebe el clúster de Amazon EMR de forma pormenorizada

Una técnica útil cuando se intenta localizar el origen de un error consiste en reiniciar el clúster y enviar los pasos al mismo uno a uno. Esto le permite comprobar los resultados de cada paso antes de procesar el siguiente y le ofrece la oportunidad de corregir y volver a ejecutar un paso que ha fallado. Esto también tiene la ventaja de cargar los datos de entrada solo una vez.

Para probar un clúster paso a paso

1. Lance un clúster nuevo, con keep alive y la protección de terminación habilitados. Keep alive mantiene el clúster en ejecución después de que haya procesado todos los pasos pendientes. La protección de terminación evita que un clúster se cierre en caso de que se produzca un error. Para obtener más información, consulte Configuración de un clúster de Amazon EMR para que continúe o finalice después de la ejecución de pasos y Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental.
2. Envíe un paso al clúster. Para obtener más información, consulte Envío del trabajo a un clúster de Amazon EMR.
3. Cuando el paso completa el procesamiento, compruebe los errores en los archivos de registro del paso. Para obtener más información, consulte Paso 4: examine los archivos de registro

de Amazon EMR. La manera más rápida de localizar estos archivos de registro consiste en conectar al nodo maestro y visualizar los archivos de registro allí. Los archivos de registro del paso no aparecen hasta que el paso se ejecuta durante algún tiempo, finaliza o genera un error.

4. Si el paso se realiza correctamente sin error, ejecutar el siguiente paso. Si se produjeron errores, investigar el error en los archivos de registro. Si se trató de un error en el código, corríjalo y vuelva a ejecutar el paso. Continúe hasta que todos los pasos se ejecuten sin errores.
5. Cuando haya terminado la depuración del clúster y desea terminarlo, tendrá que terminarlo manualmente. Esto es necesario porque el clúster se lanzó con la protección de terminación habilitada. Para obtener más información, consulte Uso de la protección de finalización para proteger sus clústeres de Amazon EMR de un cierre accidental.

Solución de problemas de un clúster de Amazon EMR lento

Esta sección describe el proceso de solución de problemas de un clúster que sigue en ejecución, pero que tarda mucho en devolver los resultados. Para obtener más información sobre qué hacer si el clúster ha terminado con un código de error, consulte Solución de problemas de un clúster de Amazon EMR que ha fallado con un código de error

Amazon EMR le permite especificar el número y el tipo de instancias en el clúster. Estas especificaciones son los medios principales que afectan a la velocidad con que la que se completa el procesamiento de datos. Una cosa que podrías considerar es volver a ejecutar el clúster, esta vez especificando EC2 instancias con más recursos o especificando un número mayor de instancias en el clúster. Para obtener más información, consulte Configuración del hardware y las redes de los clústeres de Amazon EMR.

Los siguientes temas le guiarán a través del proceso de identificación de causas alternativas de un clúster lento.

Temas

- Paso 1: recopile datos sobre el problema con el clúster de Amazon EMR
- Paso 2: compruebe el entorno del clúster de Amazon EMR
- Paso 3: examine los archivos de registro del clúster de Amazon EMR
- Paso 4: compruebe el estado de la instancia y el clúster de Amazon EMR
- Paso 5: comprobar si hay grupos suspendidos
- Paso 6: revise los ajustes de configuración del clúster de Amazon EMR

- Paso 7: examine los datos de entrada del clúster de Amazon EMR

Paso 1: recopile datos sobre el problema con el clúster de Amazon EMR

El primer paso para solucionar problemas de un clúster es recopilar información sobre lo que ha fallado y el estado y la configuración actuales del clúster. Esta información se utilizará en los siguientes pasos para confirmar o descartar las posibles causas del problema.

Definir el problema

Una definición clara del problema es el primer punto de partida. Algunas preguntas que debe hacerse:

- ¿Qué esperaba que sucediera? ¿Qué pasó en su lugar?
- ¿Cuándo se produjo este problema por primera vez? ¿Con qué frecuencia ha ocurrido desde entonces?
- ¿Ha cambiado algo en la forma en que configuro o ejecuto mi clúster?

Detalles de clúster

Los siguientes detalles del clúster son útiles para ayudar a detectar problemas. Para obtener más información sobre cómo recopilar esta información, consulte Visualización del estado y los detalles del clúster de Amazon EMR.

- El identificador del clúster. (También se denomina identificador de flujo de tareas).
- Región de AWS y la zona de disponibilidad en la que se lanzó el clúster.
- El estado del clúster, incluidos los detalles del último cambio de estado.
- Tipo y número de EC2 instancias especificados para los nodos principal, principal y de tareas.

Paso 2: compruebe el entorno del clúster de Amazon EMR

Compruebe su entorno para ver si hay interrupciones en el servicio o si ha superado un límite de AWS servicio.

Temas

- Comprobar las interrupciones de servicio

- Comprobar los límites de uso
- Comprobar la configuración de subredes de Amazon VPC
- Reiniciar el clúster

Comprobar las interrupciones de servicio

Amazon EMR utiliza varios servicios de Amazon Web Services internamente. Ejecuta servidores virtuales en Amazon EC2, almacena datos y scripts en Amazon S3 e informa de las métricas a CloudWatch. Los eventos que interrumpen estos servicios son poco frecuentes, pero, cuando se producen, pueden provocar problemas en Amazon EMR.

Antes de continuar, compruebe el Panel de estado del servicio. Compruebe la región en la que lanzó el clúster para ver si hay interrupciones en alguno de estos servicios.

Comprobar los límites de uso

Si va a lanzar un clúster grande, ha lanzado varios clústeres simultáneamente o es un usuario que comparte uno Cuenta de AWS con otros usuarios, es posible que el clúster haya fallado porque ha superado el límite de AWS servicio.

Amazon EC2 limita el número de instancias de servidores virtuales que se ejecutan en una sola AWS región a 20 instancias reservadas o bajo demanda. Si lanza un clúster con más de 20 nodos o lanza un clúster que hace que el número total de EC2 instancias activas en la Cuenta de AWS suya supere las 20, el clúster no podrá lanzar todas las EC2 instancias que necesita y podría fallar. Cuando esto ocurre, Amazon EMR devuelve un error `EC2 QUOTA EXCEEDED`. Para solicitar que se AWS aumente el número de EC2 instancias que puede ejecutar en su cuenta, envíe una solicitud de aumento del límite de EC2 instancias de Amazon.

Otro factor que puede provocar que supere los límites de uso es el retraso que transcurre entre la finalización de un clúster y el momento en que libera todos sus recursos. En función de las diferencias de configuración, un clúster puede tardar entre 5 y 20 minutos terminar por completo y liberar los recursos asignados. Si aparece un error `EC2 QUOTA EXCEEDED` al intentar lanzar un clúster, puede deberse a que aún no se hayan liberado los recursos de un clúster terminado recientemente. En este caso, puedes solicitar que se aumente tu EC2 cuota de Amazon o esperar veinte minutos y volver a lanzar el clúster.

Amazon S3 limita el número de buckets creados en una cuenta a 100. Si el clúster crea un bucket nuevo que supera este límite, se producirá un error en la creación del bucket y es posible que el clúster falle.

Comprobar la configuración de subredes de Amazon VPC

Si el clúster se lanzó en una subred de Amazon VPC, la subred debe configurarse como se describe en Configuración de redes en una VPC para Amazon EMR. Además, compruebe que la subred en la que lanza el clúster tenga suficientes direcciones IP elásticas libres para asignar una a cada nodo del clúster.

Reiniciar el clúster

La ralentización de procesamiento puede deberse a una condición transitoria. Plantéese terminar y reiniciar el clúster para ver si el rendimiento mejora.

Paso 3: examine los archivos de registro del clúster de Amazon EMR

El siguiente paso consiste en examinar los archivos de registro para encontrar un código de error u otro indicio del problema que ha sufrido el clúster. Para obtener información sobre los archivos de registro disponibles, dónde encontrarlos y cómo verlos, consulte Visualización de los archivos de registro de Amazon EMR.

Es posible que sea necesario un poco de trabajo de investigación para determinar qué pasó. Hadoop ejecuta los trabajos en los intentos de tareas en varios nodos del clúster. Amazon EMR puede iniciar intentos de tareas especulativos y terminar los demás intentos de tareas que no se completen primero. Esto genera una actividad importante que se registra en los archivos de registro del controlador, stderr y syslog a medida que se produce. Además, se ejecutan varios intentos de tareas simultáneamente, pero un archivo de registro solo puede mostrar los resultados de forma lineal.

Para comenzar, compruebe los registros de acciones de arranque para ver si hay errores o cambios de configuración inesperados durante el lanzamiento del clúster. A partir de ahí, consulte los registros de pasos para identificar los trabajos de Hadoop lanzados como parte de un paso con errores. Examine los registros de trabajos de Hadoop para identificar los intentos fallidos de tareas. El registro de intentos de tarea contendrá detalles sobre la causa del error de un intento de tarea.

En las siguientes secciones, se describe cómo utilizar los distintos archivos de registro para identificar errores en el clúster.

Comprobar los registros de acción de arranque

Las acciones de arranque ejecutan scripts en el clúster a medida que se lanza. Por lo general, se utilizan para instalar software adicional en el clúster o para modificar los valores predeterminados de

los valores de configuración. La comprobación de estos registros puede proporcionar información sobre los errores que se produjeron durante la configuración del clúster, así como sobre los cambios en los ajustes de configuración que podrían afectar al rendimiento.

Comprobar los registros de pasos

Hay cuatro tipos de registros de pasos.

- controlador: contiene archivos generados por Amazon EMR (Amazon EMR) que se deben a errores encontrados al intentar ejecutar el paso. Si se produce un error en el paso durante la carga, puede encontrar el registro de seguimiento de la pila en este registro. Aquí se describen con frecuencia los errores al cargar la aplicación o al acceder a ella, así como los errores que faltan en el archivo de asignación.
- stderr: contiene los mensajes de error que se produjeron al procesar el paso. Los errores de carga de la aplicación se describen a menudo aquí. En ocasiones, este registro contiene un seguimiento de pila.
- stdout: contiene el estado generado por los ejecutables de asignación y reducción. Los errores de carga de la aplicación se describen a menudo aquí. En ocasiones, este registro contiene mensajes de error de la aplicación.
- syslog: contiene registros de software ajeno a Amazon, como Apache y Hadoop. Los errores de streaming suelen describirse aquí.

Compruebe si hay errores obvios en stderr. Si stderr muestra una lista corta de errores, el paso se detuvo rápidamente y se produjo un error. En la mayoría de los casos, esto se debe a un error en las aplicaciones de asignación y reducción que se ejecutan en el clúster.

Examine las últimas líneas del controlador y de syslog en busca de avisos de errores. Siga cualquier aviso sobre tareas con errores, especialmente si dice “Trabajo con errores”.

Comprobar los registros de intento de tarea

Si el análisis anterior de los registros de pasos reveló una o más tareas fallidas, investigue los registros de los intentos de tareas correspondientes para obtener información de error más detallada.

Comprobar los registros de daemon de Hadoop

En raras ocasiones, Hadoop podría fallar. Para comprobar si ese es el caso, consulte los registros de Hadoop. Están ubicados en `/var/log/hadoop/` en cada nodo.

Puedes usar los JobTracker registros para asignar un intento de tarea fallido al nodo en el que se ejecutó. Una vez que sepas el nodo asociado al intento de tarea, puedes comprobar el estado de la EC2 instancia que aloja ese nodo para ver si se ha producido algún problema, por ejemplo, si se ha quedado sin CPU o memoria.

Paso 4: compruebe el estado de la instancia y el clúster de Amazon EMR

Un clúster de Amazon EMR está formado por nodos que se ejecutan en instancias de Amazon EC2 . Si dichas instancias se ven limitadas por los recursos (como, por ejemplo, quedarse sin CPU o memoria), tienen problemas de conectividad de red o se terminan, la velocidad de procesamiento del clúster se resiente.

Existen hasta tres tipos de nodos en un clúster:

- nodo maestro: administra el clúster. Si experimenta problemas de rendimiento, se ve afectado todo el clúster.
- nodos principales: procesan tareas de map-reduce y mantienen Hadoop Distributed FileSystem (HDFS). Si uno de estos nodos experimenta un problema de rendimiento, puede ralentizar las operaciones de HDFS, así como el procesamiento de MapReduce. Puede añadir más nodos secundarios a un clúster para mejorar el rendimiento, pero no puede eliminar nodos secundarios. Para obtener más información, consulte Cambio manual del tamaño de un clúster de Amazon EMR en ejecución.
- nodos de tareas: procesan tareas map-reduce. Se trata exclusivamente de recursos informáticos y no almacenan datos. Puede añadir nodos de tareas a un clúster para acelerar el rendimiento o eliminar los nodos de tareas que no sean necesarios. Para obtener más información, consulte Cambio manual del tamaño de un clúster de Amazon EMR en ejecución.

Al examinar el estado de un clúster, debe examinar tanto el rendimiento global del clúster, así como el rendimiento de instancias concretas. Existen varias herramientas que puede utilizar:

Compruebe el estado del clúster con CloudWatch

Cada clúster de Amazon EMR informa de las métricas a. CloudWatch Estas métricas proporcionan información sobre el rendimiento de resumen acerca del clúster, como la carga total, utilización de HDFS, ejecución de tareas, tareas restante, bloques corruptos, etc. Al analizar las CloudWatch métricas, obtendrá una visión general de lo que está sucediendo con su clúster y puede proporcionar información sobre las causas de la ralentización del procesamiento. Además de usarlo CloudWatch para analizar un problema de rendimiento existente, puede configurar alarmas que emitan alertas

si se produce un problema de rendimiento en el futuro. CloudWatch Para obtener más información, consulte Supervisión de las métricas de Amazon EMR con CloudWatch.

Comprobar el estado del trabajo y de HDFS

Utilice la pestaña Interfaces de usuario de aplicaciones de la página de detalles del clúster para ver los detalles de las aplicaciones de YARN. Para determinadas aplicaciones, puede consultar información adicional y tener acceso a los logs directamente. Esto resulta especialmente útil para las aplicaciones Spark. Para obtener más información, consulte Visualización del historial de aplicaciones de Amazon EMR.

Hadoop proporciona una serie de interfaces web que puede utilizar para ver información. Para obtener más información sobre cómo acceder a estas interfaces web, consulte Ver las interfaces web alojadas en clústeres de Amazon EMR.

- JobTracker — proporciona información sobre el progreso del trabajo que está procesando el clúster. Puede utilizar esta interfaz para identificar si se ha bloqueado un trabajo.
- HDFS NameNode : proporciona información sobre el porcentaje de uso de HDFS y el espacio disponible en cada nodo. Puede utilizar esta interfaz para identificar cuando HDFS se ve limitado por los recursos y requiere capacidad adicional.
- TaskTracker — proporciona información sobre las tareas del trabajo que está procesando el clúster. Puede utilizar esta interfaz para identificar cuando se ha bloqueado una tarea.

Comprueba el estado de la instancia con Amazon EC2

Otra forma de buscar información sobre el estado de las instancias de tu clúster es usar la EC2 consola de Amazon. Como cada nodo del clúster se ejecuta en una EC2 instancia, puedes usar las herramientas proporcionadas por Amazon EC2 para comprobar su estado. Para obtener más información, consulte Ver instancias de clúster en Amazon EC2.

Paso 5: comprobar si hay grupos suspendidos

Un grupo de instancias queda suspendido cuando encuentra demasiados errores al intentar lanzar nodos. Por ejemplo, si nodos nuevos devuelven error repetidamente al llevar a cabo acciones de arranque, el grupo de instancias, después de un tiempo, pasará al estado `SUSPENDED` en lugar intentar de forma continua aprovisionar nuevos nodos.

Un nodo podría no cargarse si:

- Hadoop o el clúster están estropeados por algún motivo y no aceptan un nuevo nodo en el clúster
- Una acción de arranque falla en el nuevo nodo
- El nodo no funciona correctamente y no puede iniciar sesión con Hadoop

Si un grupo de instancias está en estado `SUSPENDED` y el clúster está en estado `WAITING`, puede añadir un paso de clúster para restablecer el número deseado de nodos secundarios y de tareas. Al añadir el paso se reanuda el procesamiento del clúster y coloca el grupo de instancias de nuevo en estado `RUNNING`.

Para obtener más información sobre cómo restablecer un clúster en estado suspendido, consulte Estado de suspensión.

Paso 6: revise los ajustes de configuración del clúster de Amazon EMR

Los ajustes de configuración especifican detalles acerca de cómo se ejecuta un clúster, como cuántas veces se vuelve a intentar una tarea y la cantidad de memoria que hay disponible para clasificación. Al lanzar un clúster con Amazon EMR, existen ajustes específicos de Amazon EMR además de los ajustes de configuración estándar de Hadoop. Los ajustes de configuración se almacenan en el nodo principal del clúster. Puede comprobar los ajustes de configuración para asegurarse de que su clúster tenga los recursos que necesita para ejecutarse de forma eficaz.

Amazon EMR define los ajustes de configuración de Hadoop predeterminados que utiliza para lanzar un clúster. Los valores se basan en la AMI y el tipo de instancia que especifique para el clúster. Puede modificar los ajustes de configuración a partir de los valores predeterminados mediante una acción de arranque o especificando nuevos valores en parámetros de ejecución de trabajo. Para obtener más información, consulte Creación de acciones de arranque para instalar software adicional con un clúster de Amazon EMR. Para determinar si una acción de arranque ha cambiado los ajustes de configuración, compruebe los registros de la acción de arranque.

Amazon EMR registra los ajustes de Hadoop utilizados para ejecutar cada trabajo. Los datos de registro se almacenan en un archivo `job_`*`job-id`*`_conf.xml` denominado en el `/mnt/var/log/hadoop/history/` directorio del nodo principal, donde *`job-id`* se sustituyen por el identificador del trabajo. Si ha activado el archivado de registros, estos datos se copian en Amazon S3 en la `logs/`*`date`*`/`*`jobflow-id`*`/jobs` carpeta, donde aparece la fecha en que *`date`* se ejecutó el trabajo y *`jobflow-id`* es el identificador del clúster.

Los siguientes ajustes de configuración de trabajo de Hadoop son especialmente útiles para investigar los problemas de rendimiento. Para obtener más información acerca de los ajustes

de configuración de Hadoop y cómo afectan al comportamiento de Hadoop, visite http://hadoop.apache.org/docs/.

> **Warning**
>
> 1. Establecer `dfs.replication` en 1 en clústeres con menos de cuatro nodos puede conllevar la pérdida de datos del HDFS si un solo nodo deja de funcionar. Se recomienda que utilice un clúster con al menos cuatro nodos principales para las cargas de trabajo de producción.
> 2. Amazon EMR no permitirá que los clústeres escalen los nodos principales por debajo de `dfs.replication`. Por ejemplo, si `dfs.replication = 2`, el número mínimo de nodos principales es 2.
> 3. Cuando utiliza el escalado administrado, el escalado automático o decide cambiar el tamaño del clúster manualmente, se recomienda que establezca `dfs.replication` en 2 o más.

Opción de configuración	Descripción
dfs.replication	El número de nodos de HDFS en los que un bloque único (como el bloque de disco duro) se copian para producir un entorno similar a RAID. Determina el número de nodos de HDFS que contienen una copia del bloque.
io.sort.mb	Memoria total disponible para clasificación. Este valor debería ser 10x io.sort.factor. Este ajuste también puede utilizarse para calcular la memoria total utilizada por el nodo de tareas calculando io.sort.mb multiplicado por mapred.tasktracker.ap.tasks.maximum.
io.sort.spill.percent	Utilizado durante la clasificación, momento en que el disco empezará a utilizarse porque la memoria de clasificación asignada se está llenando.
mapred.child.java.opts	Obsoleto. Utilice mapred.map.child.java.opts y mapred.reduce.child.java.opts en su lugar. Las opciones de Java que se TaskTracker utilizan al lanzar una JVM para

Opción de configuración	Descripción
	ejecutar una tarea en ella. Un parámetro común es “-Xmx” para configurar el tamaño de memoria máximo.
mapred.map.child.java.opts	Las opciones de Java se TaskTracker utilizan al lanzar una JVM para ejecutar una tarea de mapa en ella. Un parámetro común es “-Xmx” para configurar el tamaño de montón de memoria máximo.
mapred.map.tasks.speculative.execution	Determina si los intentos de tarea Map de la misma tarea se pueden lanzar en paralelo.
mapred.reduce.tasks.speculative.execution	Determina si los intentos de tarea Reduce de la misma tarea se pueden lanzar en paralelo.
mapred.map.max.attempts	El número máximo de veces que se puede intentar una tarea Map. Si todos fallan, entonces la tarea Map se marca como error.
mapred.reduce.child.java.opts	Las opciones de Java que se TaskTracker utilizan al lanzar una JVM reducen la cantidad de tareas que se pueden ejecutar en ella. Un parámetro común es “-Xmx” para configurar el tamaño de montón de memoria máximo.
mapred.reduce.max.attempts	El número máximo de veces que se puede intentar una tarea Reduce. Si todos fallan, entonces la tarea Map se marca como error.
mapred.reduce.slowstart.completed.maps	La cantidad de tareas Map que deben completar antes de intentar las tareas Reduce. Si no se espera lo suficiente, se podrían devolver errores “Demasiados errores de recuperación” en los intentos.
mapred.reuse.jvm.num.tasks	Una tarea se ejecuta dentro de una única JVM. Especifica cuántas tareas podría reutilizar la misma JVM.

Opción de configuración	Descripción
mapred.tasktracker.map.tasks.maximum	La cantidad máxima de tareas que se pueden ejecutar en paralelo por nodo de tareas durante el mapeo.
mapred.tasktracker.reduce.tasks.maximum	La cantidad máxima de tareas que se pueden ejecutar en paralelo por nodo de tareas durante la reducción.

Si las tareas de clúster utilizan mucha memoria, puede mejorar el rendimiento utilizando menos tareas por nodo secundario y reduciendo el tamaño de montón de rastreador de trabajos.

Paso 7: examine los datos de entrada del clúster de Amazon EMR

Compruebe los datos de entrada. ¿Están distribuidos de manera uniforme entre los valores clave? Si los datos están muy sesgados hacia uno o varios valores clave, la carga de procesamiento podría estar asignada a un pequeño número de nodos, mientras que los demás nodos están inactivos. Esta distribución desequilibrada de trabajo puede dar lugar a tiempos de procesamiento más lentos.

Un ejemplo de conjunto de datos desequilibrado sería la ejecución de un clúster para alfabetizar palabras, pero disponer de un conjunto de datos que contenga solo palabras que comienzan con la letra "a". Cuando el trabajo se ha planificado, los valores de procesamiento del nodo que comienzan por "a" serían abrumadores, mientras que los nodos que procesan palabras que comienzan por otras letras estarían inactivos.

Solución de problemas comunes al utilizar Amazon EMR AWS con Lake Formation

Esta sección le guía por el proceso de resolución de problemas comunes al utilizar Amazon EMR con AWS Lake Formation.

No se permite el acceso al lago de datos

Debe optar explícitamente por el filtrado de datos en los clústeres de Amazon EMR antes de poder analizar y procesar los datos de su lago de datos. Cuando se produzca un error en el acceso a los datos, verá un mensaje `Access is not allowed` genérico en el resultado de las entradas de su bloc de notas.

Para activar y permitir el filtrado de datos en Amazon EMR, consulte Permitir el filtrado de datos en Amazon EMR en la Guía para desarrolladores de AWS Lake Formation para obtener instrucciones.

Vencimiento de la sesión

El tiempo de espera para EMR Notebooks y Zeppelin se controla mediante el rol de IAM para la configuración `Maximum CLI/API session duration` de Lake Formation. El valor predeterminado para esta configuración es de una hora. Cuando se agote el tiempo de espera de una sesión, verá el siguiente mensaje en el resultado de las entradas de su bloc de notas al intentar ejecutar los comandos de Spark SQL.

```
Error 401    HTTP ERROR: 401 Problem accessing /sessions/2/statements.
Reason:  JWT token included in request failed validation.
Powered by Jetty:// 9.3.24.v20180605
 org.springframework.web.client.HttpClientErrorException: 401 JWT token included in
 request failed validation…
```

Para validar su sesión actualice la página. Se le pedirá que vuelva a autenticarse con su IdP y se le redirigirá de vuelta al bloc de notas. Puede continuar ejecutando consultas después de la segunda autenticación.

No hay permisos para el usuario en la tabla solicitada

Al intentar acceder a una tabla a la que no tiene acceso, verá la siguiente excepción en el resultado de las entradas de su bloc de notas al intentar ejecutar comandos de Spark SQL.

```
org.apache.spark.sql.AnalysisException:
 org.apache.hadoop.hive.ql.metadata.HiveException: Unable to fetch table table.
Resource does not exist or requester is not authorized to access requested
 permissions.
(Service: AWSGlue; Status Code: 400; Error Code: AccessDeniedException; Request ID: …
```

Para obtener acceso a la tabla, debe conceder acceso al usuario. Para ello, actualice los permisos asociados a esta tabla en Lake Formation.

Consulta de datos entre cuentas compartidos con Lake Formation

Al usar Amazon EMR para acceder a los datos compartidos desde otra cuenta, algunas bibliotecas de Spark intentarán llamar a la operación de la API `Glue:GetUserDefinedFunctions`. Como

las versiones 1 y 2 de los permisos AWS RAM administrados no admiten esta acción, recibirá el siguiente mensaje de error:

```
"ERROR: User: arn:aws:sts::012345678901:assumed-role/my-
spark-role/i-06ab8c2b59299508a is not authorized to perform:
glue:GetUserDefinedFunctions on resource: arn:exampleCatalogResource
because no resource-based policy allows the glue:GetUserDefinedFunctions
action"
```

Para resolver este error, el administrador del lago de datos que creó el recurso compartido debe actualizar los permisos AWS RAM administrados asociados al recurso compartido. La versión 3 de los permisos administrados de AWS RAM permite a las entidades principales llevar a cabo la acción `glue:GetUserDefinedFunctions`.

Si crea un nuevo recurso compartido, Lake Formation aplica la última versión del permiso AWS RAM gestionado de forma predeterminada y no es necesario que realice ninguna acción. Para habilitar el acceso a los datos entre cuentas para los recursos compartidos existentes, debe actualizar los permisos AWS RAM administrados a la versión 3.

Puede ver los AWS RAM permisos asignados a los recursos compartidos con usted en AWS RAM. Los siguientes permisos se incluyen en la versión 3:

```
Databases
  AWSRAMPermissionGlueDatabaseReadWriteForCatalog
  AWSRAMPermissionGlueDatabaseReadWrite

Tables
  AWSRAMPermissionGlueTableReadWriteForCatalog
  AWSRAMPermissionGlueTableReadWriteForDatabase

AllTables
  AWSRAMPermissionGlueAllTablesReadWriteForCatalog
  AWSRAMPermissionGlueAllTablesReadWriteForDatabase
```

Para actualizar la versión de los permisos AWS RAM administrados de los recursos compartidos existentes

Usted (administrador del lago de datos) puede actualizar los permisos AWS RAM administrados a una versión más reciente siguiendo las instrucciones de la Guía del AWS RAM usuario o puede

revocar todos los permisos existentes para el tipo de recurso y volver a concederlos. Si revoca los permisos, AWS RAM elimina el AWS RAM recurso compartido asociado al tipo de recurso. Al volver a conceder los permisos, AWS RAM crea nuevos recursos compartidos adjuntando la última versión de los permisos gestionados. AWS RAM

Inserción, creación y alteración de tablas

No se admite la inserción, creación ni alteración de tablas en las bases de datos protegidas por políticas de Lake Formation. Al realizar estas operaciones, verá la siguiente excepción en el resultado de las entradas de su bloc de notas al intentar ejecutar los comandos de Spark SQL:

```
java.io.IOException:
 com.amazon.ws.emr.hadoop.fs.shaded.com.amazonaws.services.s3.model.AmazonS3Exception:
           Access Denied (Service: Amazon S3; Status Code: 403; Error Code:
 AccessDenied; Request ID: …
```

Para obtener más información, consulte [Limitaciones de la integración de Amazon EMR con. AWS Lake Formation]

Escritura de aplicaciones que lanzan y administran clústeres de Amazon EMR

Puede acceder a la funcionalidad proporcionada por la API de Amazon EMR llamando a las funciones de contenedor en una de las. AWS SDKs AWS SDKs Proporcionan funciones específicas del idioma que integran la API del servicio web y simplifican la conexión al servicio web, ya que gestionan muchos de los detalles de la conexión por usted. Para obtener más información sobre cómo llamar a Amazon EMR mediante uno de los SDKs, consulte. Se usa SDKs para llamar a Amazon EMR APIs

Temas

- End-to-end Ejemplo de código fuente Java de Amazon EMR
- Conceptos comunes para las llamadas a la API de Amazon EMR
- Se usa SDKs para llamar a Amazon EMR APIs
- Administrar las Service Quotas de Amazon EMR

> **Important**
>
> La máxima velocidad de solicitudes para Amazon EMR; es una solicitud cada diez segundos.

End-to-end Ejemplo de código fuente Java de Amazon EMR

Los desarrolladores pueden llamar a la API de Amazon EMR mediante código Java personalizado para hacer lo mismo que con la consola de Amazon EMR o la CLI. En esta sección se proporcionan los end-to-end pasos necesarios para instalar AWS Toolkit for Eclipse y ejecutar un ejemplo de código fuente de Java completamente funcional que añade pasos a un clúster de Amazon EMR.

> **Note**
>
> Este ejemplo se centra en Java, pero Amazon EMR también admite varios lenguajes de programación con una colección de Amazon EMR. SDKs Para obtener más información, consulte Se usa SDKs para llamar a Amazon EMR APIs.

Este código fuente de Java de ejemplo muestra cómo realizar las siguientes tareas a través de la API de Amazon EMR:

- Recupere AWS las credenciales y envíelas a Amazon EMR para realizar llamadas a la API
- Configurar un nuevo paso personalizado y un nuevo paso predefinido
- Agregar nuevos pasos a un clúster de Amazon EMR existente
- Recupere un paso IDs de clúster de un clúster en ejecución

> **Note**
>
> Este ejemplo muestra cómo añadir pasos a un clúster existente y requiere que tenga un clúster activo en su cuenta.

Antes de comenzar, instale una versión del Eclipse IDE for Java EE Developers (IDE de Eclipse para Java EE Developers) que coincida con su plataforma informática. Para obtener más información, consulte Eclipse Downloads.

A continuación, instale el complemento Database Development para Eclipse.

Para instalar el complemento Database Development para Eclipse

1. Abra el IDE de Eclipse.
2. Elija Help (Ayuda) e Install New Software (Instalar software nuevo).
3. En el campo Work with (Trabajar con), escriba **`http://download.eclipse.org/releases/kepler`** o la ruta que coincida con el número de versión de su IDE de Eclipse.
4. En la lista de elementos, elija Database Development (Desarrollo de bases de datos) y Finish (Finalizar).
5. Reinicie Eclipse cuando se le solicite.

A continuación, instale el Kit de herramientas para Eclipse para hacer que estén disponibles las útiles plantillas de proyecto de código fuente preconfiguradas.

Para instalar el Kit de herramientas para Eclipse

1. Abra el IDE de Eclipse.

2. Elija Help (Ayuda) e Install New Software (Instalar software nuevo).
3. En el campo Work with (Trabajar con), escriba **`https://aws.amazon.com/eclipse`**.
4. En la lista de elementos, seleccione AWS Toolkit for Eclipse y Terminar.
5. Reinicie Eclipse cuando se le solicite.

A continuación, cree un nuevo proyecto de AWS Java y ejecute el código fuente de Java de muestra.

Para crear un nuevo proyecto de AWS Java

1. Abra el IDE de Eclipse.
2. Elija File (Archivo), New (Nuevo) y Other (Otros).
3. En el cuadro de diálogo Seleccionar un asistente, seleccione Proyecto de Java de AWS y Siguiente.
4. En el cuadro de diálogo Nuevo proyecto AWS Java, introduzca en el **`Project name:`** campo el nombre del nuevo proyecto, por ejemplo**`EMR-sample-code`**.
5. Elija Configurar AWS cuentas..., introduzca sus claves de acceso públicas y privadas y elija Finalizar. Para obtener más información sobre cómo crear las claves de acceso, consulte ¿Cómo obtengo credenciales de seguridad? en la Referencia general de Amazon Web Services.

   > **Note**
   >
   > No debe incrustar las claves de acceso directamente en el código. El SDK de Amazon EMR le permite colocar las claves de acceso en ubicaciones conocidas para que no tenga que mantenerlas en el código.

6. En el proyecto de Java nuevo, haga clic con el botón derecho en la carpeta src y, a continuación, elija New (Crear) y Class (Clase).
7. En el cuadro de diálogo Java Class (Clase de Java), en el campo Name (Nombre), introduzca un nombre para la clase nueva, por ejemplo **`main`**.
8. En la sección Which method stubs would you like to create? (¿Qué stubs de método le gustaría crear?), elija public static void main(String[] args) y Finish (Finalizar).
9. Escriba el código fuente de Java dentro de la clase nueva y añada las instrucciones import adecuadas para las clases y los métodos del ejemplo. Para su comodidad se muestra a continuación el código fuente completo.

> **Note**
>
> En el siguiente código de ejemplo, sustituya el ID de clúster de ejemplo (JobFlowId)*j-xxxxxxxxxxxx*, por un ID de clúster válido en su cuenta que se encuentre en el comando siguiente AWS Management Console o mediante el siguiente AWS CLI comando:
>
> ```
> aws emr list-clusters --active | grep "Id"
> ```
>
> Además, sustituya la ruta de Amazon S3 de ejemplo, *s3://path/to/my/jarfolder*, por la ruta válida a su archivo JAR. Por último, sustituya el nombre de clase de ejemplo *com.my.Main1* por el nombre correcto de la clase en su JAR, si procede.

```
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();

  // Run a bash script using a predefined step in the StepFactory helper class
  StepFactory stepFactory = new StepFactory();
  StepConfig runBashScript = new StepConfig()
    .withName("Run a bash script")
    .withHadoopJarStep(stepFactory.newScriptRunnerStep("s3://jeffgoll/emr-scripts/
create_users.sh"))
    .withActionOnFailure("CONTINUE");

  // Run a custom jar file as a step
  HadoopJarStepConfig hadoopConfig1 = new HadoopJarStepConfig()
    .withJar("s3://path/to/my/jarfolder") // replace with the location of the jar
 to run as a step
    .withMainClass("com.my.Main1") // optional main class, this can be omitted if
 jar above has a manifest
    .withArgs("--verbose"); // optional list of arguments to pass to the jar
  StepConfig myCustomJarStep = new StepConfig("RunHadoopJar", hadoopConfig1);

  AddJobFlowStepsResult result = emr.addJobFlowSteps(new AddJobFlowStepsRequest()
    .withJobFlowId("j-xxxxxxxxxxxx") // replace with cluster id to run the steps
    .withSteps(runBashScript, myCustomJarStep));

  System.out.println(result.getStepIds());

 }
}
```

10. Elija Run (Ejecutar), Run As (Ejecutar como) y Java Application (Aplicación de Java).
11. Si el ejemplo se ejecuta correctamente, aparecerá una lista de IDs los nuevos pasos en la ventana de la consola del IDE de Eclipse. El resultado correcto es similar al siguiente:

```
[s-39BLQZRJB2E5E, s-1L6A4ZU2SAURC]
```

Conceptos comunes para las llamadas a la API de Amazon EMR

Al escribir una aplicación que llame a la API de Amazon EMR, existen varios conceptos que se aplican a la hora de llamar a una de las funciones contenedoras de un SDK.

Temas

- Puntos de conexión para Amazon EMR
- Especificar parámetros de clúster en Amazon EMR
- Zonas de disponibilidad en Amazon EMR
- Cómo utilizar archivos y bibliotecas adicionales en clústeres de Amazon EMR

Puntos de conexión para Amazon EMR

Un punto de enlace es una URL que es el punto de entrada de un servicio web. Cada solicitud de servicio web debe contener un punto de enlace. El punto final especifica la AWS región en la que se crean, describen o terminan los clústeres. Tiene el formulario `elasticmapreduce.`*`regionname`*`.amazonaws.com`. Si especifica el punto de conexión general (`elasticmapreduce.amazonaws.com`), Amazon EMR dirige la solicitud a un punto de conexión de la región predeterminada. Para las cuentas creadas el 8 de marzo de 2013 o después de esa fecha, la región predeterminada es us-west-2; en el caso de cuentas más antiguas, la región predeterminada es us-east-1.

Para obtener más información acerca de los puntos de conexión de Amazon EMR, consulte Regiones y puntos de conexión en la Referencia general de Amazon Web Services.

Especificar parámetros de clúster en Amazon EMR

Los `Instances` parámetros permiten configurar el tipo y la cantidad de EC2 instancias para crear nodos que procesen los datos. Hadoop reparte el procesamiento de los datos entre varios nodos del clúster. El nodo principal es responsable de realizar un seguimiento del estado de los nodos secundarios y de tareas y de sondear los nodos para conocer el estado del resultado de los trabajos. Los nodos secundarios y de tareas hacen el procesamiento real de los datos. Si tiene un clúster de un solo nodo, el nodo sirve tanto como nodo principal y nodo secundario.

El parámetro `KeepJobAlive` en una solicitud `RunJobFlow` determina si se debe terminar el clúster cuando se queda sin pasos de clúster que ejecutar. Defina este valor en `False` cuando sepa que el clúster se ejecuta según lo previsto. Al resolver problemas del flujo de trabajo y añadir pasos mientras la ejecución del clúster se suspende, defina el valor en `True`. Esto reduce la cantidad de tiempo y los gastos necesarios para cargar los resultados en Amazon Simple Storage Service (Amazon S3), solo para repetir el proceso después de modificar un paso para reiniciar el clúster.

Si `KeepJobAlive` es `true` así, después de conseguir que el clúster complete correctamente su trabajo, debe enviar una `TerminateJobFlows` solicitud o el clúster seguirá funcionando y generará AWS cargos.

Para obtener más información sobre los parámetros exclusivos de`RunJobFlow`, consulte RunJobFlow. Para obtener más información acerca de los parámetros genéricos en la solicitud, consulte Parámetros de solicitud comunes.

Zonas de disponibilidad en Amazon EMR

Amazon EMR utiliza las EC2 instancias como nodos para procesar los clústeres. Estas EC2 instancias tienen ubicaciones compuestas por zonas y regiones de disponibilidad. Las regiones son ubicaciones dispersas emplazadas en zonas geográficas distintas. Las zonas de disponibilidad son ubicaciones diferentes dentro de una región aisladas en caso de error en otras zonas de disponibilidad. Cada zona de disponibilidad proporciona conectividad de red económica y de baja latencia con otras zonas de disponibilidad dentro de la misma región. Para ver una lista de las regiones y los puntos de conexión de Amazon EMR, consulte Regiones y puntos de conexión en la Referencia general de Amazon Web Services.

El parámetro `AvailabilityZone` especifica la ubicación del clúster. Este parámetro es opcional y, en general, no se aconseja su uso. Cuando no se especifica `AvailabilityZone`, Amazon EMR elige automáticamente el mejor valor de `AvailabilityZone` para el clúster. Puede encontrar este parámetro útil si desea coubicar sus instancias con otras instancias en ejecución existentes y su clúster necesita leer o escribir datos de dichas instancias. Para obtener más información, consulta la Guía del EC2 usuario de Amazon.

Cómo utilizar archivos y bibliotecas adicionales en clústeres de Amazon EMR

Hay ocasiones en las que le podría interesar utilizar archivos adicionales o bibliotecas personalizadas con las aplicaciones de mapeador o reductor. Por ejemplo, podría utilizar una biblioteca que convierta un archivo PDF en texto sin formato.

Para almacenar en caché un archivo que utilice el mapeador o s reductor al utilizar Hadoop Streaming

- En el campo `args` del JAR:, añada el siguiente argumento:

  ```
  -cacheFile s3://bucket/path_to_executable#local_path
  ```

  El archivo, `local_path`, está en el directorio de trabajo del mapeador, que podría hacer referencia al archivo.

Se usa SDKs para llamar a Amazon EMR APIs

Temas

- Uso de AWS SDK para Java para crear un clúster de Amazon EMR

AWS SDKs Proporcionan funciones que integran la API y se ocupan de muchos de los detalles de la conexión, como el cálculo de las firmas, la gestión de los reintentos de las solicitudes y la gestión de los errores. SDKs También incluyen ejemplos de código, tutoriales y otros recursos que te ayudarán a empezar a escribir aplicaciones que AWS llamen. Llamar a las funciones contenedoras de un SDK puede simplificar considerablemente el proceso de escritura de una AWS aplicación.

Para obtener más información sobre cómo descargar y utilizar el AWS SDKs, consulte la SDKs sección Herramientas para Amazon Web Services.

Uso de AWS SDK para Java para crear un clúster de Amazon EMR

AWS SDK para Java Incluye tres paquetes con la funcionalidad Amazon EMR:

- com.amazonaws.services.elasticmapreduce
- com.amazonaws.services.elasticmapreduce.model
- com.amazonaws.services.elasticmapreduce.util

Para obtener más información sobre estos paquetes, consulte la referencia de la API de AWS SDK para Java.

El siguiente ejemplo ilustra cómo SDKs se puede simplificar la programación con Amazon EMR. El código de muestra siguiente utiliza el objeto `StepFactory`, una clase helper para crear los tipos de pasos de Amazon EMR comunes, para crear un clúster de Hive interactivo con la depuración habilitada.

```
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) {
  AWSCredentialsProvider profile = null;
  try {
   credentials_profile = new ProfileCredentialsProvider("default"); // specifies any
 named profile in
                                    // .aws/credentials as the credentials provider
  } catch (Exception e) {
   throw new AmazonClientException(
     "Cannot load credentials from .aws/credentials file. " +
       "Make sure that the credentials file exists and that the profile name is defined
 within it.",
     e);
  }

  // create an EMR client using the credentials and region specified in order to
  // create the cluster
  AmazonElasticMapReduce emr = AmazonElasticMapReduceClientBuilder.standard()
    .withCredentials(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());

  // specify applications to be installed and configured when EMR creates the
  // cluster
  Application hive = new Application().withName("Hive");
  Application spark = new Application().withName("Spark");
  Application ganglia = new Application().withName("Ganglia");
  Application zeppelin = new Application().withName("Zeppelin");

  // create the cluster
  RunJobFlowRequest request = new RunJobFlowRequest()
    .withName("MyClusterCreatedFromJava")
    .withReleaseLabel("emr-5.20.0") // specifies the EMR release version label, we
 recommend the latest release
    .withSteps(enabledebugging)
```

```
    .withApplications(hive, spark, ganglia, zeppelin)
    .withLogUri("s3://path/to/my/emr/logs") // a URI in S3 for log files is required
 when debugging is enabled
    .withServiceRole("EMR_DefaultRole") // replace the default with a custom IAM
 service role if one is used
    .withJobFlowRole("EMR_EC2_DefaultRole") // replace the default with a custom EMR
 role for the EC2 instance
                         // profile if one is used
    .withInstances(new JobFlowInstancesConfig()
      .withEc2SubnetId("subnet-12ab34c56")
      .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());

 }

}
```

Como mínimo, debe pasar un rol de servicio y un rol de flujo de trabajo correspondientes a EMR_ DefaultRole y EC2 EMR_ _, respectivamente. DefaultRole Puede hacerlo invocando este comando para la misma cuenta. AWS CLI En primer lugar, vea si ya existen los roles:

```
aws iam list-roles | grep EMR
```

Se mostrarán tanto el perfil de la instancia (EC2EMR_ _DefaultRole) como el rol de servicio (EMR_DefaultRole), si existen:

```
"RoleName": "EMR_DefaultRole",
            "Arn": "arn:aws:iam::AccountID:role/EMR_DefaultRole"
            "RoleName": "EMR_EC2_DefaultRole",
            "Arn": "arn:aws:iam::AccountID:role/EMR_EC2_DefaultRole"
```

Si los roles predeterminados no existen, puede utilizar el siguiente comando para crearlos:

```
aws emr create-default-roles
```

Administrar las Service Quotas de Amazon EMR

Temas

- ¿Qué son las Service Quotas de Amazon EMR?
- Cómo administrar Service Quotas para Amazon EMR
- Cuándo configurar los eventos de EMR en CloudWatch

Los temas de esta sección describen las cuotas de servicio de EMR (anteriormente denominadas límites de servicio), cómo gestionarlas y cuándo es ventajoso utilizar CloudWatch eventos en lugar de cuotas de servicio para supervisar los clústeres y activar acciones. AWS Management Console

¿Qué son las Service Quotas de Amazon EMR?

Su AWS cuenta tiene cuotas de servicio predeterminadas, también conocidas como límites, para cada AWS servicio. El servicio de EMR tiene dos tipos de límites:

- Límites de recursos: puede usar EMR para crear EC2 recursos. Sin embargo, estos EC2 recursos están sujetos a cuotas de servicio. Las limitaciones de recursos de esta categoría son las siguientes:
  - El número máximo de clústeres activos que se pueden ejecutar a la vez.
  - El número máximo de instancias activas por grupo de instancias.
- Límites APIs activados: cuando se utiliza el EMR APIs, los dos tipos de limitaciones son:
  - Límite de ráfaga: es el número máximo de llamadas a la API que se pueden realizar a la vez. Por ejemplo, el número máximo de solicitudes de AddInstanceFleet API que puedes realizar por segundo está establecido en 5 de calls/second forma predeterminada. Esto implica que el límite de ráfagas de llamadas a la AddInstanceFleet API es de 5 llamadas por segundo o que, en un momento dado, puedes realizar un máximo de 5 llamadas a la AddInstanceFleet API. Sin embargo, después de usar el límite de ráfaga, las llamadas subsiguientes estarán limitadas por el límite de frecuencia.
  - Límite de frecuencia: es la tasa de reabastecimiento de la capacidad de ampliación de la API. Por ejemplo, la tasa de reabastecimiento de AddInstanceFleet llamadas se establece en 0,5 de calls/second forma predeterminada. Esto significa que cuando alcance el límite de ráfaga, tendrás que esperar al menos 2 segundos (0,5 llamadas/segundo x 2 segundos = 1 llamada) para realizar la llamada a la API. Si realiza una llamada antes, el servicio web de EMR le limitará. En cualquier momento, solo puede realizar tantas llamadas como tenga la capacidad de

ampliación sin que se le limite. Cada segundo adicional que espere, la capacidad de ampliación aumentará en 0,5 llamadas hasta alcanzar el límite máximo de 5, que es el límite de ráfaga.

Cómo administrar Service Quotas para Amazon EMR

Service Quotas es una AWS función que puede utilizar para ver y gestionar sus cuotas o límites de servicio de Amazon EMR desde una ubicación central, mediante la AWS Management Console API o la. AWS CLI Para obtener más información sobre la consulta de las cuotas y la solicitud de aumentos, consulte AWS Service Quotas en la Referencia general de Amazon Web Services.

> **Important**
>
> Las cuotas de servicio se pueden aplicar de forma independiente y, en ocasiones, sus efectos pueden superponerse. Algunas tienen un alcance específico APIs y otras imponen límites a nivel de cuenta. En concreto, se aplican las siguientes cuotas a nivel de cuenta:
>
> - La velocidad máxima a la que se repone la cubeta para todas las operaciones de EMR
> - La cantidad máxima de solicitudes de API que puede realizar por segundo
>
> Ten en cuenta que si solicitas correctamente un aumento de cuota específico y sigues viendo restricciones (por ejemplo, en el caso de las solicitudes a la API), es posible que una de estas cuotas a nivel de cuenta siga limitando las tasas de llamadas a la API. Para solucionar problemas, puedes usar la consola de cuotas del servicio, que está disponible en https://console.aws.amazon.com/servicequotas/casa, para comprobar los límites de las cuotas y solicitar aumentos. Si solicitas aumentos y sigues viendo restricciones, ponte en contacto con el servicio de asistencia. También puede encontrar los límites y descripciones predeterminados para las cuotas del servicio de EMR en la guía de referencia AWS general.

Para algunos APIs, organizar un CloudWatch evento puede ser una mejor opción que aumentar las cuotas de servicio. También puedes ahorrar tiempo configurando alarmas y activando el aumento de solicitudes de forma proactiva antes de alcanzar la cuota de servicio. CloudWatch Para obtener más información, consulta Cuándo configurar los eventos de EMR en CloudWatch.

Cuándo configurar los eventos de EMR en CloudWatch

Para algunos sondeos APIs, como DescribeCluster, y DescribeStep ListClusters, configurar un CloudWatch evento puede reducir el tiempo de respuesta a los cambios y liberar las cuotas de servicio. Por ejemplo, si tiene una función de Lambda configurada para ejecutarse cuando el estado de un clúster cambie, por ejemplo, cuando se complete un paso o se termine un clúster, puede usar ese activador para iniciar la siguiente acción del flujo de trabajo en lugar de esperar al siguiente sondeo. De lo contrario, si tiene EC2 instancias de Amazon dedicadas o funciones de Lambda que consultan constantemente la API de EMR para detectar cambios, no solo desperdiciará recursos informáticos, sino que también podría alcanzar su cuota de servicio.

Los siguientes son algunos casos en los que podría beneficiarse de la transición a una arquitectura basada en eventos.

Caso 1: Sondear EMR mediante llamadas a la DescribeCluster API para completar los pasos

Example El sondeo de EMR mediante DescribeCluster API requiere completar los pasos

Un patrón habitual consiste en enviar un paso a un clúster en ejecución y sondear Amazon EMR para conocer el estado del paso, normalmente mediante la DescribeCluster tecla o. DescribeStep APIs Esta tarea también se puede realizar con un retraso mínimo conectándose al evento Step Status Change de Amazon EMR.

Este evento contiene la siguiente información en su carga útil.

```
{
  "version": "0",
  "id": "999cccaa-eaaa-0000-1111-123456789012",
  "detail-type": "EMR Step Status Change",
  "source": "aws.emr",
  "account": "123456789012",
  "time": "2016-12-16T20:53:09Z",
  "region": "us-east-1",
  "resources": [],
  "detail": {
    "severity": "ERROR",
    "actionOnFailure": "CONTINUE",
    "stepId": "s-ZYXWVUTSRQPON",
    "name": "CustomJAR",
```

```
    "clusterId": "j-123456789ABCD",
    "state": "FAILED",
    "message": "Step s-ZYXWVUTSRQPON (CustomJAR) in Amazon EMR cluster j-123456789ABCD
 (Development Cluster) failed at 2016-12-16 20:53 UTC."
  }
}
```

En el mapa de detalles, una función de Lambda podría analizar “state”, “stepId” o “clusterId” para encontrar la información pertinente.

Caso 2: Sondeo de EMR en busca de clústeres disponibles para ejecutar flujos de trabajo

Example Sondeo de EMR en busca de clústeres disponibles para ejecutar flujos de trabajo

Un patrón de los clientes que ejecutan varios clústeres es ejecutar los flujos de trabajo en los clústeres tan pronto como estén disponibles. Si hay muchos clústeres en ejecución y es necesario realizar un flujo de trabajo en un clúster que está en espera, un patrón podría consistir en sondear los EMR mediante DescribeCluster llamadas a la ListClusters API para los clústeres disponibles. Otra forma de reducir el retraso a la hora de saber cuándo un clúster está listo para un paso sería procesar el evento State Change de Amazon EMR.

Este evento contiene la siguiente información en su carga útil.

```
{
  "version": "0",
  "id": "999cccaa-eaaa-0000-1111-123456789012",
  "detail-type": "EMR Cluster State Change",
  "source": "aws.emr",
  "account": "123456789012",
  "time": "2016-12-16T20:43:05Z",
  "region": "us-east-1",
  "resources": [],
  "detail": {
    "severity": "INFO",
    "stateChangeReason": "{\"code\":\"\"}",
    "name": "Development Cluster",
    "clusterId": "j-123456789ABCD",
    "state": "WAITING",
    "message": "Amazon EMR cluster j-123456789ABCD ..."
  }
```

```
}
```

Para este caso, se podría configurar una función de Lambda para enviar inmediatamente un flujo de trabajo en espera a un clúster en cuanto su estado cambie a EN ESPERA.

Caso 3: Sondeo de EMR para la terminación del clúster

Example Sondeo de EMR para la terminación del clúster

Un patrón común de clientes que ejecutan muchos clústeres de EMR consiste en sondear Amazon EMR en busca de los clústeres terminados para que ya no se les envíe trabajo. Puede implementar este patrón con las llamadas a la DescribeCluster ListClusters API o mediante el evento Amazon EMR Cluster State Change en.

Tras la terminación del clúster, el evento emitido se parece al siguiente ejemplo.

```
{
  "version": "0",
  "id": "1234abb0-f87e-1234-b7b6-000000123456",
  "detail-type": "EMR Cluster State Change",
  "source": "aws.emr",
  "account": "123456789012",
  "time": "2016-12-16T21:00:23Z",
  "region": "us-east-1",
  "resources": [],
  "detail": {
    "severity": "INFO",
    "stateChangeReason": "{\"code\":\"USER_REQUEST\",\"message\":\"Terminated by user
 request\"}",
    "name": "Development Cluster",
    "clusterId": "j-123456789ABCD",
    "state": "TERMINATED",
    "message": "Amazon EMR Cluster jj-123456789ABCD (Development Cluster) has
 terminated at 2016-12-16 21:00 UTC with a reason of USER_REQUEST."
  }
}
```

La sección “detalles” de la carga incluye el id. de clúster y el estado sobre los que se puede actuar.

AWS Glosario

Para obtener la AWS terminología más reciente, consulte el AWS glosario de la Glosario de AWS Referencia.

Historial de documentos

En la siguiente tabla se describen los cambios notables en la Guía de administración de Amazon EMR desde la última versión de Amazon EMR. Para obtener más información sobre versiones publicadas específicas de Amazon EMR, consulte Acerca de las versiones de Amazon EMR.

Cambio	Descripción	Fecha de lanzamiento de la nueva versión
Actualización de la política administrada	Actualizaciones de Amazon EMR de las políticas AWS gestionadas — Actualizaciones de Amazon EMR de las políticas AWS gestionadas — Permisos adicionales para Amazon Policy_v2. EMRService	4 de marzo de 2025
Versión inicial	Publicación inicial de la guía de administración de Amazon EMR para las versiones 4.x y posteriores de Amazon EMR. Para obtener más información sobre las versiones anteriores de Amazon EMR, consulte la Guía para desarrolladores de Amazon EMR.	24 de julio de 2015

Las traducciones son generadas a través de traducción automática. En caso de conflicto entre la traducción y la version original de inglés, prevalecerá la version en inglés.